

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:15:31 ; Search time 186 Seconds
(without alignments)
1322.861 Million cell updates/sec

Title: US-09-853-880A-17
Perfect score: 3026
Sequence: 1 MECLYFLGFLLLAARLPD.....PPGNQEKDPLLKNQEFKGV 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Genesep1980s.*
- 2: Genesep1990s.*
- 3: Genesep2000s.*
- 4: Genesep2001s.*
- 5: Genesep2002s.*
- 6: Genesep2003as.*
- 7: Genesep2003bs.*
- 8: Genesep2004s.*
- 9: Genesep2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	2 AAW35382	AAW35382 Murine me
2	3026	100.0	560	3 AAB11329	Aab11329 Human lun
3	3026	100.0	560	5 ABB78200	Abb78200 Amino aci
4	3026	100.0	560	5 ABB74961	Abb74961 Human lun
5	3026	100.0	560	5 ABP61881	Abp61881 Human lun
6	3026	100.0	560	6 ABUS6592	Abus6592 Lung canc
7	3026	100.0	560	6 ABG72962	Abg72962 Human oet
8	3026	100.0	560	6 ABU70852	Abu70852 Human adi
9	3026	100.0	560	7 ADA28315	Ada28315 Human lun
10	3026	100.0	560	7 ADG95620	Adg95620 Human NOV
11	3026	100.0	560	7 ADH36879	Adh36879 Human lun
12	3026	100.0	560	7 ADJ68660	Adj68660 Human hea
13	3026	100.0	560	7 ADL14995	Adl14995 Human NMB
14	3026	100.0	560	7 ADN39940	Adn39940 Cancer/an
15	3026	100.0	560	8 ADH56342	Adh56342 Human nmb
16	3026	100.0	560	8 ADJ75569	Adj75569 Marker ge
17	3026	100.0	560	8 ADM56682	Adm56682 Human lun
18	3026	100.0	560	8 ADQ18310	Adq18310 Human sof
19	3026	100.0	560	8 ADP23127	Adp23127 PRO polytp
20	3026	100.0	560	9 ADU98374	Adu98374 Lung tumo
21	3026	100.0	560	9 AEB10182	Aeb10182 Cancer re
22	3026	100.0	561	8 ADU06546	Adu06546 Novel bro
23	3026	100.0	563	8 ADR66659	Adr66659 Human pro
24	3026	100.0	563	8 ADR66317	Adr66317 Human pro

25	3026	100.0	563	8 ADR66339	Adr66339 Human pro
26	3026	100.0	563	8 ADR66681	Adr66681 Human pro
27	3007	99.4	572	7 ADD78235	Add78235 Human CGD
28	3006	99.3	572	5 AAU83612	Aau83612 Human PRO
29	3006	99.3	572	5 ADY31782	Ady31782 Novel hum
30	3006	99.3	572	6 ABU80759	Abu80759 Human PRO
31	3006	99.3	572	6 ABO33725	Ab033725 Novel hum
32	3006	99.3	572	6 ABU82068	Abu82068 Novel hum
33	3006	99.3	572	6 ABJ72248	Abj72248 Human PRO
34	3006	99.3	572	6 ABJ72376	Abj72376 Human PRO
35	3006	99.3	572	6 ABO34271	Ab034271 Human sec
36	3006	99.3	572	7 ABJ72078	Abj72078 Human mem
37	3006	99.3	572	7 ADB83532	Adb83532 Novel hum
38	3006	99.3	572	7 ADB80638	Adb80638 Novel hum
39	3006	99.3	572	7 ADB73179	Adb73179 Novel hum
40	3006	99.3	572	7 ADB78261	Adb78261 Novel hum
41	3006	99.3	572	7 ADB84909	Adb84909 Human PRO
42	3006	99.3	572	7 ADB78015	Adb78015 Novel hum
43	3006	99.3	572	7 ADB87081	Adb87081 Human PRO
44	3006	99.3	572	7 ADB84663	Adb84663 Human PRO
45	3006	99.3	572	7 ADB83778	Adb83778 Novel hum

ALIGNMENTS

RESULT 1

AAW35382
ID AAW35382 standard; protein; 560 AA.
XX
AC AAW35382;
XX
DT 26-FEB-1998 (first entry)
XX
DE Murine metastatic nucleic acid sequence product.
XX
KW Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;
KW treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening.
XX
OS Mus musculus.
XX
PN WO9718454-A2. ✓
XX
PD 22-MAY-1997.
XX
PF 15-NOV-1996; 96WO-US018567.
XX
PR 16-NOV-1995; 95US-0006838P.
PR 30-JAN-1996; 96US-00594031.
XX
(THOM/) THOMPSON T.
XX
Thompson T;
XX
WPI; 1997-289397/26.
XX
Identifying tumour metastatic sequences - by introducing transfected
cells into host mammal and analysing primary and metastatic sequences by
differential display PCR.
XX
Disclosure; Fig 12CI; 102pp; English.
XX
Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse
embryos. The UGS cells were infected with retroviruses, cultured and
implanted under the renal capsule of mice. Reconstitutions were harvested
5 weeks later, when they showed signs of distress from the tumour burden.
Metastatised tumours were isolated from a site outside the renal capsule.
RNA was isolated from primary tumours and metastases, reverse transcribed
and subjected to differential display PCR. The sequences were analysed to
obtain metastatic sequences, e.g. the sequence encoding the present
sequence. The method can be used to detect, diagnose and treat disorders
related to metastasis, or treat malignant or non-malignant disorders, can
e.g. hyperplasia, dysplasia and hypertrophy. The metastatic sequence, can

CC be used to screen a biological sample for metastasis, and it or its
CC expression product may also be used to treat a metastatic disorder
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVNGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVNGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Qy 61 VKKGDNRWKNKSGGRVQAVLTSDSPALVGSNITFAVNLIIFRCQKEDANGNIYVEKNC 120
Db 61 VKKGDNRWKNKSGGRVQAVLTSDSPALVGSNITFAVNLIIFRCQKEDANGNIYVEKNC 120
Qy 121 RNEAGLSADPYVYNTWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNTWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180
Qy 181 GOYFQKLGRCRVSrvsvntantvlgpqlmevtyvrrhgrayvpiqkvdyvvtvtdqiipfv 240
Db 181 GOYFQKLGRCRVSrvsvntantvlgpqlmevtyvrrhgrayvpiqkvdyvvtvtdqiipfv 240
Qy 241 TMFQKNDNRNSDDETFKDLPIIMFDVLIHDPESHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIIMFDVLIHDPESHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLELSRIPDENQCIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLELSRIPDENQCIN 360
Qy 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPPESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPPESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Qy 421 CEITQNTVCSrPVDVDEMCLLTVRTFTNGSGTYCVNLTLDGDDTSLALTSLISVDPDRPAS 480
Db 421 CEITQNTVCSrPVDVDEMCLLTVRTFTNGSGTYCVNLTLDGDDTSLALTSLISVDPDRPAS 480
Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560

RESULT 2
AAB11329
ID AAB11329 standard; protein; 560 AA.
XX
AC AAB11329;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human lung cancer-associated protein L528S.
XX
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection.
XX
OS Homo sapiens.
XX
FN WO200061612-A2. ✓
XX
PD 19-OCT-2000.
XX
PF 03-APR-2000; 2000WO-US008896.
XX
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.

PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480584.
PR 22-FEB-2000; 2000US-00510376.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
XX WPI; 2000-628399/60.
DR
XX
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX protein is used for detecting and monitoring progression of lung cancer
XX in a patient.
XX
XX Claim 3; Page 203-204; 261pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer
XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVNGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVNGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Qy 61 VKKGDNRWKNKSGGRVQAVLTSDSPALVGSNITFAVNLIIFRCQKEDANGNIYVEKNC 120
Db 61 VKKGDNRWKNKSGGRVQAVLTSDSPALVGSNITFAVNLIIFRCQKEDANGNIYVEKNC 120
Qy 121 RNEAGLSADPYVYNTWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNTWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180
Qy 181 GOYFQKLGRCRVSrvsvntantvlgpqlmevtyvrrhgrayvpiqkvdyvvtvtdqiipfv 240
Db 181 GOYFQKLGRCRVSrvsvntantvlgpqlmevtyvrrhgrayvpiqkvdyvvtvtdqiipfv 240
Qy 241 TMFQKNDNRNSDDETFKDLPIIMFDVLIHDPESHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIIMFDVLIHDPESHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLELSRIPDENQCIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLELSRIPDENQCIN 360
Qy 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPPESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPPESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Qy 421 CEITQNTVCSrPVDVDEMCLLTVRTFTNGSGTYCVNLTLDGDDTSLALTSLISVDPDRPAS 480
Db 421 CEITQNTVCSrPVDVDEMCLLTVRTFTNGSGTYCVNLTLDGDDTSLALTSLISVDPDRPAS 480
Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560

```

RESULT 3
ABB78200
ID ABB78200 standard; protein; 560 AA.
XX
AC ABB78200;
XX
DT 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of human HGFIN.
XX
KW Human; cell differentiation; white blood cell; bone marrow cell;
KW haematopoietic growth factor inducible neurokin-1; HGFIN;
KW progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;
KW acute lymphocytic leukemia; chronic myeloid leukemia;
KW chronic lymphocytic leukemia; Hodgkin's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..485
FT /note= "extracellular domain"
FT Misc-difference 256
FT /note= "Leu encoded by CC"
FT Domain 485..508
FT /note= "transmembrane domain"
FT Domain 509..560
FT /note= "intracellular domain"
XX
PN WO200262947-A2.
XX
PD 15-AUG-2002.
XX
PF 20-OCT-2001; 2001WO-US050204.
XX
PR 20-OCT-2000; 2000US-0241881P.
XX
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI Pranela R;
XX
DR WPI; 2002-657531/70.
DR N-PSDB; ABQ78551.
XX
XX Hematopoietic growth factor inducible neurokin-1 type polypeptide and
PT polynucleotide for treating a disease associated with abnormal bone
PT marrow cell differentiation or proliferation, e.g. leukemia.
XX
PS Claim 12; Page 123-125; 125pp; English.
XX
CC The present sequence represents haematopoietic growth factor inducible
CC neurokin-1 type (HGFIN) polypeptide. HGFIN induces white blood cell
CC differentiation and inhibits progenitor proliferation. HGFIN polypeptides
CC and polynucleotides are useful for treating a disease associated with
CC abnormal bone marrow cell differentiation or proliferation, especially
CC acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid
CC leukemia, chronic lymphocytic leukemia, Hodgkin's and non-Hodgkin's
CC disease
XX
SQ Sequence 560 AA;
Query Match 100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYFFLGFLLAARLPDAAKRFHDVLGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFFLGFLLAARLPDAAKRFHDVLGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120

```

```
XX PS Disclosure; Page 281-283; 374pp; English.
XX CC The present invention describes human lung tumour proteins. Human lung
XX CC tumour proteins and polynucleotides have cytostatic and immunostimulant
XX CC activities, and can be used in vaccine production. Compositions
XX CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations, or antigen presenting cells that express
XX CC the lung tumour proteins are useful for treating lung cancer or
XX CC stimulating an immune response. ABL48959 to ABL49300 and ABL74946 to
XX CC ABL75070 represent sequences used in the exemplification of the present
XX CC invention
XX SQ Sequence 560 AA;
XX
XX Query Match 100.0%; Score 3026; DB 5; Length 560;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-262;
XX Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
QY 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
QY 181 QYFQKLGRCRSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
DB 181 QYFQKLGRCRSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
QY 241 TMFQKNDNRNSDETFLKDLPIMFDLIHPDPSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKDLPIMFDLIHPDPSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
DB 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
QY 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
DB 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
DB 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
DB 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
DB 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
QY 481 PLRMANSALISVGCLAIPTVVISLIVYKHKYENPIENSPGNVYRSKLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCLAIPTVVISLIVYKHKYENPIENSPGNVYRSKLSVFLNRAKAVF 540
QY 541 PPGNQEKDPLLNQBFKGV 560
DB 541 PPGNQEKDPLLNQBFKGV 560
XX
XX RESULT 5
XX ID ABP61881
XX AC ABP61881 standard; protein; 560 AA.
XX
XX AC ABP61881;
XX
XX DT 07-OCT-2002 (first entry)
XX
XX DE Human lung cancer associated protein sequence SEQ ID NO:225.
XX
XX KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
XX OS Homo sapiens.
```

```
XX WO200247534-A2. ✓
XX PD 20-JUN-2002.
XX
XX 30-NOV-2001; 2001WO-US047576.
XX
XX 12-DEC-2000; 2000US-00735705.
XX 07-MAY-2001; 2001US-00850716.
XX 28-JUN-2001; 2001US-00897778.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;
XX Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2002-583465/62.
XX
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
XX the polynucleotides, useful in pharmaceutical compositions such as
XX vaccines and as markers to indicate the presence of lung cancer.
XX
XX Example 1; Page 289-290; 381pp; English.
XX
XX The present invention describes isolated human lung carcinoma
XX polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
XX activity, and can be used in gene therapy and in vaccines. Compositions
XX comprising (I) or (II) can be used for stimulating an immune response in
XX a patient and for treating lung cancer in a patient. Oligonucleotides of
XX (I) can be used for detecting the presence of a cancer in a patient, by
XX obtaining a biological sample from the patient, contacting the biological
XX sample with the oligonucleotide, detecting in the sample an amount of
XX polynucleotide that hybridises to the oligonucleotide and comparing the
XX amount of polynucleotide that hybridises to the oligonucleotide to a
XX predetermined cut-off value, and determining the presence of a cancer in
XX the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
XX vaccines. (I) is useful as a marker to indicate the presence or absence
XX of a cancer such as lung cancer. ABQ92145 to ABQ92496 and ABP61866 to
XX ABP61992 represent sequences used in the exemplification of the present
XX invention
XX
XX Sequence 560 AA;
XX
XX Query Match 100.0%; Score 3026; DB 5; Length 560;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-262;
XX Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGPFPHPGWRWNFIYVFTL 180
QY 181 QYFQKLGRCRSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
DB 181 QYFQKLGRCRSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
QY 241 TMFQKNDNRNSDETFLKDLPIMFDLIHPDPSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKDLPIMFDLIHPDPSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
DB 301 HTYVLNGTFSNLNLTAKAAGPCPPPPPPRPSKPTSLGPNLPSLRIIPDENCOIN 360
QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
DB 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
```

Db 361 RYGFQATITIVGILEVNIQMTDVLMPVPWPESSLDIFVVTCCQSGIPEVCTIISDPT 420
 QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDPPAS 480
 Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDPPAS 480
 QY 481 PLRMANSALISVGCLAI FVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVP 540
 Db 481 PLRMANSALISVGCLAI FVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVP 540
 QY 541 PFGNQEKDPLLNQKQFPGVS 560
 Db 541 PFGNQEKDPLLNQKQFPGVS 560

RESULT 6
 ABUS6592
 ID ABUS6592 standard; protein; 560 AA.
 XX AC
 XX ABUS6592;
 DT 02-APR-2003 (first entry)
 XX DE Lung cancer-associated polypeptide #185.
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 XX Unidentified.
 OS WO200286443-A2.
 PN 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012476.
 XX 18-APR-2001; 2001US-0284770P.
 XX 10-MAY-2001; 2001US-0290492P.
 XX 09-NOV-2001; 2001US-0339245P.
 XX 13-NOV-2001; 2001US-0350666P.
 XX 29-NOV-2001; 2001US-0334370P.
 XX 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR N-PSDB; ABX76321.
 XX

PT Detecting a lung cancer-associated transcript in a cell from a patient
 for treating lung cancer, by contacting a biological sample from the
 patient with a polynucleotide that exhibits increased or decreased
 expression in lung cancer.
 XX
 XX Claim 27; Page 328; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 transcript in a cell from a patient, comprising contacting a biological
 sample from the patient with a polynucleotide that selectively hybridizes
 to a sequence that is at least 80 % identical to a gene that exhibits
 increased or decreased expression in lung cancer samples. Lung cancer-
 associated polynucleotides and polypeptides are used for identifying a
 compound that modulates a lung cancer-associated polypeptide, for
 inhibiting proliferation of a lung cancer-associated cell to treat lung
 cancer in a patient and for treating a mammal having lung cancer by
 administering a modulatory compound identified. The methods are useful
 for treating lung cancer, such as small cell lung cancer, non-small cell
 lung cancer or other benign or precancerous lesions, e.g. atelectasis,

CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 6; Length 560;
 Best Local Similarity 100.0%; Pred. No. 5.7e-262;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRPDVLGNERPSAYMREHNQLNGWSSDENDWNEKLYP 60
 Db 1 MECLYYFLGFLLLAARLPDAAKRPDVLGNERPSAYMREHNQLNGWSSDENDWNEKLYP 60
 QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCKEDANGNIYVEKNC 120
 Db 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCKEDANGNIYVEKNC 120
 QY 121 RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDGPHPHGWRRWFIYVFHTL 180
 Db 121 RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDGPHPHGWRRWFIYVFHTL 180
 QY 181 GOYFQKLGRCVRSVNTANTVLPQLMETVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
 Db 181 GOYFQKLGRCVRSVNTANTVLPQLMETVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
 QY 241 TMFQKNDNRSDETFKLPIMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 Db 241 TMFQKNDNRSDETFKLPIMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPBRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
 Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPBRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
 QY 361 RYGFQATITIVGILEVNIQMTDVLMPVPWPESSLDIFVVTCCQSGIPEVCTIISDPT 420
 Db 361 RYGFQATITIVGILEVNIQMTDVLMPVPWPESSLDIFVVTCCQSGIPEVCTIISDPT 420
 QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDPPAS 480
 Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDPPAS 480
 QY 481 PLRMANSALISVGCLAI FVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVP 540
 Db 481 PLRMANSALISVGCLAI FVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVP 540
 QY 541 PFGNQEKDPLLNQKQFPGVS 560
 Db 541 PFGNQEKDPLLNQKQFPGVS 560

RESULT 7
 ABG72962

ID ABG72962 standard; protein; 560 AA.

XX AC ABG72962;

XX DT 08-APR-2003 (first entry)
 XX DE Human osteoactivin homologue, nmb.

XX KW Human; osteoactivin; osteopathic; antiinflammatory; vaccine;
 XX bone cell differentiation stimulator; gene therapy; bone formation;
 XX osteoporosis; periodontal disease; ectopic bone formation; osteopetrosis;
 XX bone disorder; osteogenesis; enzyme.
 OS Homo sapiens.
 XX US2002151486-A1.
 PN

PR	17-DEC-1999;	99US-00466396.	
PR	30-DEC-1999;	99US-00476496.	
PR	10-JAN-2000;	2000US-00480884.	
PR	23-FEB-2000;	2000US-00510376.	
PR	04-APR-2000;	2000US-00542615.	
PR	28-JUN-2000;	2000US-00606421.	
PR	02-AUG-2000;	2000US-00630940.	
PR	21-AUG-2000;	2000US-00643597.	
PR	15-SEP-2000;	2000US-00662786.	
PR	09-OCT-2000;	2000US-00685696.	
PR	12-DEC-2000;	2000US-00735705.	
PR	07-MAY-2001;	2001US-00850716.	
PR	28-JUN-2001;	2001US-00897778.	
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;		
PI	Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;		
PI	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;		
XX			
DR	WPI, 2003-540798/51.		
XX			
PT	New isolated polynucleotides and polypeptides useful for diagnosing,		
PT	preventing and/or treating cancer, particularly lung cancer.		
XX			
XX	Example 7; Page 194-195; 296pp; English.		
XX			
CC	The invention describes isolated polynucleotides and polypeptides useful		
CC	for diagnosing, preventing and/or treating cancer, particularly lung		
CC	cancer. A new isolated polynucleotide comprises: any of the 22 fully		
CC	defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the		
CC	specification; complements of the nucleotide sequences cited above; at		
CC	least 10 contiguous residues of the nucleotide sequences cited above; a		
CC	sequence that hybridise to any of the nucleotide sequences under highly		
CC	stringent conditions; a sequence that is at least 75 or 90% identical to		
CC	the above nucleotide sequences; or degenerate variants of the above		
CC	nucleotide sequences. The composition and methods are useful in		
CC	diagnosing, preventing and/or treating cancer, particularly lung cancer,		
CC	in gene therapy and in vaccines. This is the amino acid sequence encoded		
CC	by a human lung tumour cDNA isolated from a lung squamous cell carcinoma		
CC	that may be useful in the diagnosis and treatment of lung cancer and		
CC	other disorders.		
XX			
SQ	Sequence 560 AA;		
	Query Match	100.0%; Score 3026; DB 7; Length 560;	
	Best Local Similarity	100.0%; Pred. No. 5.7e-262;	
	Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Qy	1	MECLYYFLGFLLLAARLPDAAKPHDVLGNRPSAYNRHNQLNGWSDENDWNEKLYP	60
Db	1	MECLYYFLGFLLLAARLPDAAKPHDVLGNRPSAYNRHNQLNGWSDENDWNEKLYP	60
Qy	61	VWRKGDNRKNSWKGGRVQAVLTSDSPALVGSNTTFAVNLIFPRCKQKDANGNIYVERKC	120
Db	61	VWRKGDNRKNSWKGGRVQAVLTSDSPALVGSNTTFAVNLIFPRCKQKDANGNIYVERKC	120
Qy	121	RNEAGLSADPVYVNTAWSEDSDENGTGQSHNVFPDGKPPPHHGRWRNFIYVFHTL	180
Db	121	RNEAGLSADPVYVNTAWSEDSDENGTGQSHNVFPDGKPPPHHGRWRNFIYVFHTL	180
Qy	181	GQYFQKLGRCRSVRVSVNTAVTLGQPMVEVYRRHGRAYVPVIAQVKDQVYVTDQIPVFV	240
Db	181	GQYFQKLGRCRSVRVSVNTAVTLGQPMVEVYRRHGRAYVPVIAQVKDQVYVTDQIPVFV	240
Qy	241	TMFQKDRNRSDETFKDLPIFMDVLHDPGSHFLNYSNTINKWSFGDNTGLFVSTNHTVN	300
Db	241	TMFQKDRNRSDETFKDLPIFMDVLHDPGSHFLNYSNTINKWSFGDNTGLFVSTNHTVN	300
Qy	301	HTYYVLNGTFSNLNLTAKAAAPCGCPPPPPPPPSKPTPSLGAGDNPFLSRLIPDENCQIN	360
Db	301	HTYYVLNGTFSNLNLTAKAAAPCGCPPPPPPPPSKPTPSLGAGDNPFLSRLIPDENCQIN	360

```
Qy 361 RYGHFOATITIVIEGLEVNIIQMTDVLMPVPWPRESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Db 361 RYGHFOATITIVIEGLEVNIIQMTDVLMPVPWPRESSLIIDFVVTQCGSIPTVECTIIISDPT 420
Qy 421 CEITONTVCSVDVDEMCCLLVTRTFNGSGTYCVNLTGLDDTSLALSTLISVDRDPAS 480
Db 421 CEITONTVCSVDVDEMCCLLVTRTFNGSGTYCVNLTGLDDTSLALSTLISVDRDPAS 480
Qy 481 PLRWANSALISVGCCLAIPTVTISLLVYKKHKEYNPISNPGNVRSKGLSVFLNRAKAVF 540
Db 481 PLRWANSALISVGCCLAIPTVTISLLVYKKHKEYNPISNPGNVRSKGLSVFLNRAKAVF 540
Qy 541 PFGNQEKDPLLNQEFKGV 560
Db 541 PFGNQEKDPLLNQEFKGV 560

RESULT 10
ID ADE95620 standard; protein; 560 AA.
XX
AC ADE95620;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human NOVX28c protein.
XX
KW NOVX protein; biochemical stimulation; physiological stimulation;
KW cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;
KW antirheumatic; antiarthritic; antidiabetic; nephrotropic; dermatological;
KW immunosuppressive; anti-HIV; antiinflammatory; neuroprotective;
KW neutropic; antipeptidic; antiparkinsonian; antiasthmatic; neuroleptic;
KW antidepressant; anti-allergic; gynaecological; gene therapy; vaccine;
KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; hypertension;
KW cancer; obesity; rheumatoid arthritis; diabetes; glomerulonephritis;
KW psoriasis; skin disorder; AIDS; inflammation; multiple sclerosis;
KW Alzheimer's disease; Parkinson's disease; asthma; schizophrenia;
KW depression; allergy; fertility disorder; NOVX28c.
XX
OS Homo sapiens.
XX

Key Location/Qualifiers
FH Misc-difference 47
FT /label= OTHER
FT /note= "OTHER= May be substituted by Arg as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 57
FT /label= OTHER
FT /note= "OTHER= May be substituted by Glu as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 111
FT /label= OTHER
FT /note= "OTHER= May be substituted by Asp as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 130
FT /label= OTHER
FT /note= "OTHER= May be substituted by Pro as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 154
FT /label= OTHER
FT /note= "OTHER= May be substituted by Ser as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 162
FT /label= OTHER
FT /note= "OTHER= May be substituted by Leu as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 294
FT /label= OTHER
FT /note= "OTHER= May be substituted by Phe as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 324
FT /label= OTHER
FT /note= "OTHER= May be substituted by Leu as a result of a
```

```
FT Misc-difference 460
FT /label= OTHER
FT /note= "OTHER= May be substituted by Asp as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 487
FT /label= OTHER
FT /note= "OTHER= May be substituted by Ser as a result of a
FT single nucleotide polymorphism"
FT Misc-difference 526
FT /label= OTHER
FT /note= "OTHER= May be substituted by Arg as a result of a
FT single nucleotide polymorphism"
PN WO2003050245-A2.
XX
PD 19-JUN-2003.
XX
XX 03-DEC-2002; 2002WO-US038594.
XX
XX 05-DEC-2001; 2001US-0336600P.
XX 07-DEC-2001; 2001US-0338285P.
XX 12-DEC-2001; 2001US-0341346P.
XX 17-DEC-2001; 2001US-0341477P.
XX 17-DEC-2001; 2001US-0341540P.
XX 20-DEC-2001; 2001US-0342592P.
XX 27-DEC-2001; 2001US-0344297P.
XX 31-DEC-2001; 2001US-0344903P.
XX 17-APR-2002; 2002US-0373288P.
XX 15-MAY-2002; 2002US-0380981P.
XX 17-MAY-2002; 2002US-0381495P.
XX 28-MAY-2002; 2002US-0383534P.
XX 28-MAY-2002; 2002US-0383744P.
XX 29-MAY-2002; 2002US-0383829P.
XX 29-MAY-2002; 2002US-0384024P.
XX 07-AUG-2002; 2002US-0401788P.
XX 26-AUG-2002; 2002US-0406353P.
XX 31-OCT-2002; 2002US-00401788.
XX 02-DEC-2002; 2002US-00406353.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Chillakuru RA;
XX Edinger SR, Gerlach VL, Gorman L, Gould-Rothberg BE, Guo X;
XX Jeffers WE, Ji W, Li L, Malyankar UM, Miller CE, Murphey R;
XX Patturajan M, Peyman JA, Rastelli L, Rieger DK, Shenoy SG;
XX Smithson G, Starling G, Taupier RJ, Voss EZ, Zhong H, Zhong M;
XX WPI; 2003-513974/48.
XX N-PSDB; ADE95619.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX
XX Claim 2; SEQ ID NO 152; 211pp; English.
XX
XX This invention relates to novel NOVX proteins, and the DNA sequence which
XX encode them, having properties related to stimulation of biochemical or
XX physiological responses in a cell, a tissue, an organ or an organism.
XX Compounds which modulate the proteins of the invention may have cardiac,
XX antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic,
XX antiarthritic, antidiabetic, nephrotropic, dermatological,
XX immunosuppressive, anti-HIV, antiinflammatory, neuroprotective,
XX neutropic, antipeptidic, antiparkinsonian, antiasthmatic, neuroleptic,
XX antidepressant, anti-allergic or gynaecological activities. The DNA
XX sequences of the invention may be useful for gene therapy whilst the
XX protein sequences may allow the development of a vaccine. The protein is
XX useful in the manufacture of a medicament for treating a syndrome
XX associated with a human disease. The invention may be useful in
XX diagnosing, treating or preventing NOVX-associated disorders, for example
XX cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,
```


CC rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin
 CC disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease,
 CC Parkinson's disease, asthma, schizophrenia, depression, allergies or
 CC fertility disorders. The nucleic acids may further be used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The present sequence is the amino acid
 CC sequence of the human NOVX28c protein of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;
 Best Local Similarity 100.0%; Pred. No. 5,7e-262;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAARFHDVLGNERPSAYMRHNLGWSDDNDWNEKLYP 60
 Db 1 MECLYFGLFLLAARLPDAARFHDVLGNERPSAYMRHNLGWSDDNDWNEKLYP 60
 QY 61 VVKRGMWKNKSGRQVAVLTSDSPALVGSNITFAVNLIPRCOKEDANGNIYVEKNC 120
 Db 61 VVKRGMWKNKSGRQVAVLTSDSPALVGSNITFAVNLIPRCOKEDANGNIYVEKNC 120
 QY 121 RNEAGLSADPYVNTAWSEDSDGNGTQSHNVFPDGKPPHPGWRWNFIYVFHTL 180
 Db 121 RNEAGLSADPYVNTAWSEDSDGNGTQSHNVFPDGKPPHPGWRWNFIYVFHTL 180
 QY 181 GOYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVYVVDQIPV 240
 Db 181 GOYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVYVVDQIPV 240
 QY 241 TFMQKDRNSDETFLKOLPIMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 Db 241 TFMQKDRNSDETFLKOLPIMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360
 Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360
 QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420
 Db 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420
 QY 421 CBITQNTVCSVDVDEMCLLTVRRTFNGSGTVCVNLTLGDDTSLATSTLISVPDRDPAS 480
 Db 421 CBITQNTVCSVDVDEMCLLTVRRTFNGSGTVCVNLTLGDDTSLATSTLISVPDRDPAS 480
 QY 481 PLRMANSAISVGCCLAFVTVISLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540
 Db 481 PLRMANSAISVGCCLAFVTVISLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540
 QY 541 PFGNQEKDPLLNQKQFKGVS 560
 Db 541 PFGNQEKDPLLNQKQFKGVS 560

RESULT 11

ADH36879

ID ADH36879 standard; protein; 560 AA.

XX

AC ADH36879;

XX

DT 11-MAR-2004 (first entry)

XX

DE Human lung cancer-related protein #16.

XX

KW lung cancer; lung tumour; immune response stimulation;

KW

tumour protein specific T cell; human.

XX

OS Homo sapiens.

XX

PN WO2003086175-A2.

XX

PD 23-OCT-2003.

XX

PF 07-APR-2003; 2003WO-US010945.

XX

PR 05-APR-2002; 2002US-00117982.

PR

04-DEC-2002; 2002US-00313986.

XX

PA (CORI-) CORIXA CORP.

XX

PI Mericle B, Fanger GR, Vedvick TS, Carter D, Watanabe Y;

PI

Henderson RA, Kalos MD, Spies GA, Foy TM, Fan L, Wang T, McNabb A;

PI

Reed SG;

XX

WPI; 2003-845247/78.

XX

Use of compositions comprising lung tumor polypeptides or

PT

polynucleotides, for inhibiting lung cancer or tumor progression,

PT

stimulating immune response, or stimulating and/or expanding T cells

PT

specific for a tumor protein.

XX

Disclosure; SEQ ID NO 225; 456pp; English.

PS

The invention comprises a method for inhibiting the development of lung

CC

cancer, the method involves the use of compositions which contain lung

CC

tumor polypeptides/polynucleotides. The methods and compositions of the

CC

invention are useful for inhibiting lung cancer development or tumour

CC

progression, stimulating immune response, or stimulating and/or expanding

CC

T cells specific for a tumour protein. The present amino acid sequence

CC

represents a human lung cancer-related protein.

XX

Sequence 560 AA;

SQ

Query Match 100.0%; Score 3026; DB 7; Length 560;
 Best Local Similarity 100.0%; Pred. No. 5,7e-262;
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAARFHDVLGNERPSAYMRHNLGWSDDNDWNEKLYP 60
 Db 1 MECLYFGLFLLAARLPDAARFHDVLGNERPSAYMRHNLGWSDDNDWNEKLYP 60
 QY 61 VVKRGMWKNKSGRQVAVLTSDSPALVGSNITFAVNLIPRCOKEDANGNIYVEKNC 120
 Db 61 VVKRGMWKNKSGRQVAVLTSDSPALVGSNITFAVNLIPRCOKEDANGNIYVEKNC 120
 QY 121 RNEAGLSADPYVNTAWSEDSDGNGTQSHNVFPDGKPPHPGWRWNFIYVFHTL 180
 Db 121 RNEAGLSADPYVNTAWSEDSDGNGTQSHNVFPDGKPPHPGWRWNFIYVFHTL 180
 QY 181 GOYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVYVVDQIPV 240
 Db 181 GOYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVYVVDQIPV 240
 QY 241 TFMQKDRNSDETFLKOLPIMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 Db 241 TFMQKDRNSDETFLKOLPIMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
 QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360
 Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360
 QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420
 Db 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420
 QY 421 CBITQNTVCSVDVDEMCLLTVRRTFNGSGTVCVNLTLGDDTSLATSTLISVPDRDPAS 480
 Db 421 CBITQNTVCSVDVDEMCLLTVRRTFNGSGTVCVNLTLGDDTSLATSTLISVPDRDPAS 480
 QY 481 PLRMANSAISVGCCLAFVTVISLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540
 Db 481 PLRMANSAISVGCCLAFVTVISLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540
 QY 541 PFGNQEKDPLLNQKQFKGVS 560
 Db 541 PFGNQEKDPLLNQKQFKGVS 560

XX Disclosure; SEQ ID NO 7; 342pp; English.

XX The invention relates to a compound comprising a binding moiety which selectively binds to a protein or polypeptide listed in the specification (e.g. human autotaxin polypeptide or human CD24 signal transducer polypeptide), and a further moiety. The compound is useful in medicine or in the treatment, imaging, diagnosis or prognosis of mantle cell lymphomas (MCL). It is used in preparing a medicament for treating MCL, a diagnostic or prognostic agent for MCL, or an agent for imaging MCL cells in the body of an individual. This sequence corresponds to one of the polypeptides of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60

QY 61 VMKRGDMRWKNSKGRVQAVLTSPPALVGSNITFAVNLIPRCOKEDANGNIYVKNC 120
DB 61 VMKRGDMRWKNSKGRVQAVLTSPPALVGSNITFAVNLIPRCOKEDANGNIYVKNC 120

QY 121 RNEAGLSADPYVYNTANVTLSQSHHNPDPGKPPHPGWRNRNFIYVFHTL 180
DB 121 RNEAGLSADPYVYNTANVTLSQSHHNPDPGKPPHPGWRNRNFIYVFHTL 180

QY 181 GQYFQKLGRCVRSVNTANVTLGQMEVTVYRRHGRAYVPIAQVKOVYVTDQIPV 240
DB 181 GQYFQKLGRCVRSVNTANVTLGQMEVTVYRRHGRAYVPIAQVKOVYVTDQIPV 240

QY 241 TFMQKNDNRSSDETFLKLPIMFDVLIHDPSPFLNYSYNTYKWSFGDNTGLFVSTNHTVN 300
DB 241 TFMQKNDNRSSDETFLKLPIMFDVLIHDPSPFLNYSYNTYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFLSLNTVKAAPGCPPPPPPPPPSKPTPSLGPAGDNPLELSRIPDENCQIN 360
DB 301 HTYVLNGTFLSLNTVKAAPGCPPPPPPPPPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPESLIDFVVTQCGSIPTVCTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPESLIDFVVTQCGSIPTVCTIISDPT 420

QY 421 CEITQNTVCSPVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSLSVPPDRPAS 480
DB 421 CEITQNTVCSPVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSLSVPPDRPAS 480

QY 481 PLRMANSALISVGCCLAIFVTVISLVYKHKYNPISNPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCCLAIFVTVISLVYKHKYNPISNPGNVRSKGLSVFLNRAKAVF 540

QY 541 PFGNQEKDPLLNQBFKGV 560
DB 541 PFGNQEKDPLLNQBFKGV 560

RESULT 14
ADN39940
ID ADN39940 standard; protein; 560 AA.

XX ADN39940;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C310.

XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;

KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.

OS Homo sapiens.

PN WO2003042661-A2.

PD 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368809P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevesi PA;
PI Mack DH, Murray R, Watson SR, Wilson KB, Zioltnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39723.

XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C310; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60

Qy 61 VKKGDNRWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120
Dd |||||
Qy 61 VKKGDNRWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120
Dd |||||
Qy 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGKPPHPGRRWNFIYVFTL 180
Dd |||||
Qy 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGKPPHPGRRWNFIYVFTL 180
Dd |||||
Qy 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVFTDQIPVF 240
Dd |||||
Qy 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVFTDQIPVF 240
Dd |||||
Qy 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300
Dd |||||
Qy 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300
Dd |||||
Qy 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360
Dd |||||
Qy 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360
Dd |||||
Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLLIDFVTCQGSIPTEVCTIISDPT 420
Dd |||||
Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLLIDFVTCQGSIPTEVCTIISDPT 420
Dd |||||
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGYCVNLTLGGDTSALTSLISVDPDRPAS 480
Dd |||||
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGYCVNLTLGGDTSALTSLISVDPDRPAS 480
Dd |||||
Qy 481 PLRMANSALISVGCLAIPTVVISLVYKKHKEYNPIENSPGNVRSKGLSVFLNRKAVF 540
Dd |||||
Qy 481 PLRMANSALISVGCLAIPTVVISLVYKKHKEYNPIENSPGNVRSKGLSVFLNRKAVF 540
Dd |||||
Qy 541 PFGNQEKDPLLNQEFKGS 560
Dd |||||
Qy 541 PFGNQEKDPLLNQEFKGS 560
Dd |||||

RESULT 15

ADH56342
ID ADH56342 standard; protein; 560 AA.

XX
AC ADH56342;

XX
DT 25-MAR-2004 (first entry)

XX
DE Human nmb protein SEQ ID NO:1.

XX
KW respiratory disease; human; nmb; antiinflammatory; antiaesthetic;
KW antiallergic; neuroprotective; immunosuppressive; antidiabetic;
KW antiarthritic; dermatological; gastrointestinal; antirheumatic; asthma;
KW hayfever; chronic bronchitis; chronic obstructive lung disease;
KW immune disorder; multiple sclerosis; Sjogren's disease;
KW insulin-resistant diabetes; rheumatoid arthritis; lupus erythematosus;
KW atopic dermatitis; irritable bowel disease.

XX
OS Homo sapiens.

XX
FN WO2004002516-A1. ✓

XX
PD 08-JAN-2004.

XX
PP 27-JUN-2003; 2003WO-JP008168.

XX
PR 28-JUN-2002; 2002JP-00190790.

XX
PA (TAKE) TAKEDA CHEM IND LTD.

XX
PI Nakanishi A, Iwashita H, Morita S, Matsumoto T, Yamasaki M;

XX
DR WPI; 2004-082883/08.

XX
DR N-PSDB; ADH56343.

PT Drug compositions containing inhibitor of human nmb protein activity for
PT treatment and prevention of respiratory diseases including asthma and
XX chronic obstructive lung disease.

PS Claim 1; SEQ ID NO 1; 104pp; Japanese.

XX The present invention describes drug compositions (I) for prevention and
CC treatment of respiratory diseases, which contain a compound or its salts
CC inhibiting the activity of human nmb protein or its partial peptides or a
CC protein of substantially similar sequence and activity, or inhibiting the
CC expression of the gene encoding this protein. Also described: (1)
CC antisense polynucleotides to all or part of the nmb gene, and drug
CC compositions containing them; (2) antibodies to all or part of nmb
CC protein, and drug compositions and diagnostic reagents containing them;
CC (3) diagnostic reagents containing polynucleotides encoding all or part
CC of nmb protein; (4) screening method for compounds for use in prevention
CC and treatment of respiratory diseases, using nmb protein or its partial
CC peptides, or using polynucleotides encoding all or part of nmb protein,
CC or using a tobacco smoking induced chronic obstructive lung disease model
CC mouse or an elastase-induced chronic obstructive lung disease model mouse
CC ; (5) kits for the screening method; (6) compounds identified by the
CC screening method; (7) a method for prevention and treatment of
CC respiratory diseases using the identified compounds; and (8) drug
CC compositions for prevention and treatment of respiratory diseases
CC containing a compound or its salts inhibiting the binding activity of
CC heparan sulfate proteoglycan. (I) have antiinflammatory, antiaesthetic,
CC antiallergic, neuroprotective, immunosuppressive, antidiabetic,
CC antiarthritic, dermatological, gastrointestinal and antirheumatic
CC activities. (I) can be used for the prevention, treatment and diagnosis
CC of respiratory diseases such as acute asthma, chronic asthma, hayfever,
CC chronic bronchitis and chronic obstructive lung disease, and other
CC diseases including immune disorders such as multiple sclerosis, Sjogren's
CC disease, insulin-resistant diabetes, rheumatoid arthritis, lupus
CC erythematosus, atopic dermatitis, and irritable bowel disease. The
CC present sequence represents human nmb, which is used in the
CC exemplification of the present invention.

XX
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 8; Length 560;

Best Local Similarity 100.0%; Pred. No. 5.7e-262;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGSSDNDNEKLYP 60

Dd |||||

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGSSDNDNEKLYP 60

Dd |||||

Qy 61 VWKRGDMRWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120

Dd |||||

Qy 61 VWKRGDMRWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120

Dd |||||

Qy 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGKPPHPGRRWNFIYVFTL 180

Dd |||||

Qy 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGKPPHPGRRWNFIYVFTL 180

Dd |||||

Qy 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVFTDQIPVF 240

Dd |||||

Qy 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVFTDQIPVF 240

Dd |||||

Qy 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300

Dd |||||

Qy 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300

Dd |||||

Qy 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

Dd |||||

Qy 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

Dd |||||

Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLLIDFVTCQGSIPTEVCTIISDPT 420

Dd |||||

Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLLIDFVTCQGSIPTEVCTIISDPT 420

Dd |||||

Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGYCVNLTLGGDTSALTSLISVDPDRPAS 480

Dd |||||

Db 421 CEITQNTVCSPVDVDEMCLITVVRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDPAS 480
Qy 481 PLRMANSALISVGCLAIFVTVISLLVYKHKHKNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIFVTVISLLVYKHKHKNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLKNQEPKGV 560
Db 541 FPGNQEKDPLLKNQEPKGV 560

Search completed: December 24, 2005, 04:31:49
Job time : 197 secs

THIS PAGE BLANK (uspto)

Result No.	Score	Query %			DB	ID	Description
		Match	Length	Length			
1	3026	100.0	560	2	I38065	gene NWE protein -	
2	659	21.8	626	2	S53871	Pmel 17 protein -	
3	622.5	20.6	662	2	I38400	melanoma-associate	
4	617.5	20.4	668	2	A41234	melanocyte-specific b	
5	459	15.2	491	2	A49179	melanoma antigen b	
6	136	4.5	926	1	A41105	protein-tyrosine-p	
7	120.5	4.0	446	2	T07907	hydroxyproline-ric	
8	114.5	3.8	1009	2	C64483	hypothetical prote	
9	111.5	3.7	457	2	I55976	dihydrolipoamide S	
10	111.5	3.7	555	2	S21766	dihydrolipoamide S	
11	109.5	3.6	348	2	AB3260	hypothetical memb	
12	108	3.6	3848	2	T17414	TipC protein - sli	
13	106.5	3.5	848	2	T23694	hypothetical prote	
14	106	3.5	1874	1	JQ0533	genome F21918	
15	105	3.5	492	2	C95521	protein F21D18.18	
16	104.5	3.5	363	2	G83306	probable periplasm	
17	104.5	3.5	624	1	I51581	transforming prote	
18	104.5	3.5	837	2	C69200	surface proteinase	
19	104.5	3.5	840	2	B69216	surface proteinase	
20	104.5	3.5	4302	2	A38971	polycystic kidney	
21	104	3.4	913	2	T18503	hypothetical prote	
22	102.5	3.4	674	2	A10922	ATP-dependent DNA	
23	102.5	3.4	921	2	A33718	retinoblastoma pro	
24	102.5	3.4	979	2	A35913	regulatory factor	
25	101.5	3.4	445	2	D81716	hypothetical prote	
26	101	3.3	588	2	T45564	hypothetical prote	
27	101	3.3	705	2	T16088	hypothetical prote	
28	101	3.3	2014	2	T21560	hypothetical prote	
29	100.5	3.3	221	2	T07176	extensin homolog -	

```
Qy 421 CEITONTVCSPVDVDEMCLLTVRRFTFNGSGTYCVNLTGLDDTSLALTSTLISVPRDPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRFTFNGSGTYCVNLTGLDDTSLALTSTLISVPRDPAS 480
Qy 481 PLRMANSALISVGCLAIPTVTISILLYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIPTVTISILLYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 PPGNQEKDPLLNQBEFKGVS 560
Db 541 PPGNQEKDPLLNQBEFKGVS 560

RESULT 2
S53871
Emel 17 protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C/Accession: S53871
R/Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamaneni, C.; Bennett, D.; Pick
Nucleic Acids Res. 23, 154-158, 1995
A/Title: Mouse silver mutation is caused by a single base insertion in the putative cyto
A/Reference number: S53871; MUID:95173358; PMID:7870580
A/Accession: S53871
A/Status: Preliminary
A/Molecule type: mRNA
A/Residues: 1-626 <KWO>
A/Cross-references: UNIPROT:Q60696; UNIPARC:UPI000002822F; GB:U14133; NID:g887940; PIDN:

Query Match 21.8%; Score 659; DB 2; Length 626;
Best Local Similarity 27.2%; Pred. No. 2.4e-43;
Matches 183; Conservative 99; Mismatches 211; Indels 180; Gaps 20;

Qy 7 FLGFLLLAARLPDAA--AKRFDVLGNRPSPAYMREHNQNLGWSDDNDWNEKLYPVWKR 64
Db 9 FLPLVLISALLAVGALGSRNQDLGVPRQLV-----TKTWNRQLYPEWT- 53

Qy 65 GDMWKNSWKGRVQAVLTSDSPALVGSNTITFAVNLIPPRCKEDANGNIYVEKN-CRNE 123
Db 54 -EVQGSNCWRGGQVSLRVINDPTLVGNASFSIALHPFGSQKVLDPGQVIANNTING 112

Qy 124 AGLSADPVYVNTWTAMSEDSDCGNGTGGSHNVFPDGKPPFPHHPGRRWNFIYVFHTLQY 183
Db 113 SQVWGQPVYV-----QEPDDA-----CVFPDGGPCSPGPKPKRSFVYVWKTGKY 159

Qy 184 FQKLGRCVRSVNTANVTILGPQLMEVTVYRRHG--RAYVPIAQKDVVYVTDQIPVFTM 242
Db 160 WQVLGGPVRSRIAIRHAKLGTHTMEVTVYHRRGSQYVPLAHASSTFTITDQVPFVSV 219

Qy 243 FQKNDRSSDETFLKDLPIMPDVLHDPSPHFLANTSTINYKWSFGDNTGLFVSTNHTVNT 302
Db 220 SOLQALDGETKHFLRNHLIIFALQLHDPGVLAEADLSYTWDFDGGTGLISRALDVHT 279

Qy 303 YVLNGTFSNLITVXAA-----
Db 280 YLESGSVTAQVVLQAAIPLVSCGSPVPGTTDGYMPTAEAPGTTSRQGTITTKVVGTTPGQ 339

Qy 323 CPPPPP-----
Db 340 MPTTOPSGTTVQMPTEVTATTBEQMLTSAVIDTTLAEVSTTEGTGTPTRPSGTTVAQ 399

Qy 335 PTPSLPGAGDNPL-----ELGRIPDE-----NCQINRYGHQFQATITIV 373
Db 400 ATTTEGPD-A-SPLPTQSSSTGSIPLLDODTTIMLVKROVPLDCLVLYRGSFSLALDIVQ 458

Qy 374 GILEVNIQTMDVLMPVPWPESSILDFVTCQSGIPTVETCVTIISDPTCEITQNTVCSPVD 433
Db 459 G-----IESAIIQAVPFSEGDAPFELTVSCQGLPKACMDISSPGCQPPAQRLCQSPV 512

Qy 434 VDEMCLLTVRRTF-NGSGTYCVNLTGLDDTSLALTSTLISVPRDPASPLRMANSALISV 492
Db 513 PSPCQLVLHQVLKGGSGTYCLNVSADANSIAVASTQLVVPFGDGG-----LGOAPLLV 567
```

```
Qy 493 GCIAIFVTVI--SLLV----YKHKKEYN-----PIENSPGNVVRSGLSVFLNRAKAVFPF 542
Db 568 GILLVLVAVVASLILIGILRSRAQFPKCHMVALTAAPASGLRARGLG----- 615

Qy 543 GNQEKDPLLNQBE 555
Db 616 ---ENSPLLSQQ 625

RESULT 3
I38400
melanoma-associated ME20 antigen (me20m) - human
N/Alternate names: melanoma antigen 25
C/Species: Homo sapiens (man)
C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C/Accession: I38400; A53668; A55753
R/Maresh, G.A.; Marken, J.S.; Neubauer, M.; Aruffo, A.; Hellstrom, I.; Hellstrom, K.; Ma
DNA Cell Biol. 13, 87-95, 1994
A/Title: Cloning and expression of the gene for the Melanoma-Associated ME20 Antigen.
A/Reference number: I38400; MUID:94235165; PMID:8179825
A/Accession: I38400
A/Status: Preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-662 <RES>
A/Cross-references: UNIPROT:P40967; UNIPARC:UPI000016A059; EMBL:U01874; NID:g494939; PID
R/Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.
J. Biol. Chem. 269, 20126-20133, 1994
A/Title: Molecular characterization of the melanocyte lineage-specific antigen gp100.
A/Reference number: A53668; MUID:94327568; PMID:7519602
A/Accession: A53668
A/Molecule type: mRNA
A/Residues: 1-592,594-662 <ADE>
A/Cross-references: UNIPARC:UPI000000C02D; GB:S73003; NID:g639589; PIDN:AACG0634.1; PID:
R/Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Y
Proc. Natl. Acad. Sci. U.S.A. 91, 6458-6462, 1994
A/Title: Identification of a human melanoma antigen recognized by tumor-infiltrating lym
A/Reference number: A55753; MUID:94294401; PMID:8022805
A/Accession: A55753
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-161,'P',163-592,594-662 <KAW>
A/Cross-references: UNIPARC:UPI000014ACC2
C/Keywords: glycoprotein

Query Match 20.6%; Score 622.5; DB 2; Length 662;
Best Local Similarity 25.5%; Pred. No. 1.8e-40;
Matches 168; Conservative 93; Mismatches 199; Indels 199; Gaps 17;

Qy 3 CLYYFLGLLIAARLPDAAK--RFHDVLGNRPSPAYMREHNQNLGWSDDNDWNEKLYP 60
Db 8 CL----LHLAVIGALLAVGATKVPENQDLGVSR-----QLR-----TKAWNRLQYP 50

Qy 61 VWKKGDMWRKNSWKGRVQAVLTSDSPALVGSNTITFAVNLIPPRCKEDANGNIYVEKVC 120
Db 51 EWT--EAQLDCWRGGQVSLKVSNDGPTLIGANASFSIALNPFPGSQKVLDPGQVITWNNT 108

Qy 121 RNEAGLSADPVYVYV---W---TAMSEDSDCGNGTGGSHNVFPDGKPPFPHHPGRRWNF 173
Db 109 -----IINGSQVWGQPVYIPOSTDDA-----CIFPDGPGPCSGSGSQKRSF 149

Qy 174 IYVFHTLQYFQKLGRCVRSVNTANVTILGPQLMEVTVYRRHG--RAYVPIAQKDVVYV 232
Db 150 VYVWKTWGQVQVILGGPVSGLSIGTGRAMLGTHTMEVTVYHRRGSRSVYVPLAHSSAFTI 209

Qy 233 TDQIPVFTVFMQKNDRSSDETFLKDLPIMPDVLHDPSPHFLANTSTINYKWSFGDNTGLF 292
Db 210 TDQVPFVSVSQLRALDGGKHKFLRNQPLTFALQLHDPGVLAEADLSYTWDFDGSSTGL 269

Qy 293 VSTNHTVNTVHTVLTGTFSLNITVKAAP---GPCPPPP----- 327
Db 270 ISRALVTHTYLEPGPVTAQVVLQALPLTSCGSPVPGTTDGHRTAEAPNTTAGQVPT 329

Qy 328 ----- 328
```


Db 330 TEVGTTCQAPTAEPSTGTTISVQPTTEVISTAPVQMPAETSTGMTPEKVPVSEVNGTTL 389
QY 328 -----PPRPSPKP----- 335
Db 390 AEMSTPEATGMPAEVSIIVLGGTTAAQVTTTETWETTARELPPEPEGPDASSIMSTES 449
QY 336 -TPSLGPAGDNPFLSLRPDE---NCQINRYGHFOATITIVGILEVNIQMTDVLMPVP 391
Db 450 ITGSLGPLDGTATLRLVKRQVPLDCVLYRYSFSTLQVQ-----IESAEILOAVP 503
QY 392 WPESLIDPVTTCQSIPEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRTE-NGSG 450
Db 504 SEGDAPFELTVCQGLPKAEWEISSPCQPPAQLCQVLPSPACQLVLHQLKGGSG 563
QY 451 TYCVNLTGDDTSLALTSLISVPRDRPASPLRMANSALISVGCCLAIFVTIVLSLVYKK 509
Db 564 TYCLNVSLADTNSLAVSTQIMPQEGAG--LQGVPLVIGILLVMAVVLASLIYR 619
RESULT 4
A41234
C:Species: Homo sapiens (man)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 30-Sep-1993
C:Accession: A41234
R:Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.;
Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991
A:Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on m
A:Reference number: A41234; MUID:92021023; PMID:1924386
A:Accession: A41234
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-668 <KWO>
A:Cross-references: UNIPARC:UPI00001796E0; GB:M77348

Query Match 20.4%; Score 617.5; DB 2; Length 668;
Best Local Similarity 25.8%; Pred. No. 4.6e-40;
Matches 171; Conservative 93; Mismatches 197; Indels 203; Gaps 19;
QY 3 CLYFLGFLLLAARLPDAAK--RPHDVLGNRPASVYMRHNLQNGWSSDENDWNEKLYP 60
Db 8 CL---LHLAVIGALLAVGATKVPKQNDWLGVSR-----QLR-----TKAWRQLYP 50
QY 61 VKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIIPRCQKEDANGNIYVERKNC 120
Db 51 EWT--EAQLDCWRGQVSLKVSNDGPTLIGANASFSIALNFPQSKVLPDQVIVWNT 108
QY 121 RNEAGLSADPYVN---W---TANSEDSGNGTGQSHNVFPDQKPPPHPGWRRWNP 173
Db 109 -----IINGSQVMGGQPVYPOETDDA-----CIFPDGPGCPSGSGWSQKRSF 149
QY 174 IVEHTLGOYFQKLGRCVRSVNTANVTLGQPMLEVTVYRHG--RAYVPIAQVKDVVV 232
Db 150 VYVWKTWQYVQWLGPGVSLGSGTGRAMLGTHMEVTVYHRRGSRVYVPLAHSSAFTI 209
QY 233 TDQIPVFTMFQKDRNSDSEFLKDLPIPMFDVLIHDPSSHFLNYSTINYKWSFGNTGLF 292
Db 210 TDQVPSVSQRLALDGNKHFLNQPLTFALQLHDPSPGLYAEADLSYTWDFGSSGTL 269
QY 293 VSTNHTVNTYVNLGTFSLNLTVKAAP-----GPCPPPP----- 327
Db 270 ISRAPVVTHTYLEPGPVTAQVVLQAAIPLTSCGSPVPFGTTDGHRTAEAPNTTAGQVPT 329
QY 328 ----- 327
Db 330 TEVGTTCQAPTAEPSTGTTISVQPTTEVISTAPVQMPAETSTGMTPEKVPVSEVNGTTL 389
QY 328 -----PPRPSPKP----- 335
Db 390 AEMSTPEATGMPAEVSIIVLGGTTAAQVTTTETWETTARELPPEPEGPDASSIMSTES 449
QY 336 -TPSLGPAGDNPFLSLRPDE---NCQINRYGHFOATITIVGILEVNIQMTDVLMPVP 391

Db 450 ITGSLGPLDGTATLRLVKRQVPLDCVLYRYSFSTLQVQ-----IESAEILOAVP 503
QY 392 WPESLIDPVTTCQSIPEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRTE-NGSG 450
Db 504 SEGDAPFELTVCQGLPKAEWEISSPCQPPAQLCQVLPSPACQLVLHQLKGGSG 563
QY 451 TYCVNLTGDDTSLALTSLISVPRDRPASPLRMANSALISVGCCLAIFVTIVLSLVYKK 509
Db 564 TYCLNVSLADTNSLAVSTQIMP--VFGILLTGOEAGLQGVPLVIGILLVMAVVLASL 621
QY 506 VYKK 509
Db 622 IYR 625
RESULT 5
A49179
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A49179; I45861
R:Kim, R.Y.; Wistow, G.J.
Exp. Eye Res. 55, 657-662, 1992
A:Title: The cDNA RPE1 and monoclonal antibody HMB-50 define gene products preferentiall
A:Reference number: A49179; MUID:93122163; PMID:1478275
A:Accession: A49179
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-491 <KIM>
A:Cross-references: UNIPROT:Q06154; UNIPARC:UPI0000131C85
A:Experimental source: retinal pigment epithelium
A:Note: sequence extracted from NCBI backbone (NCBIN:122438, NCBI:122439)
C:Genetics:
A:Gene: RPE1

Query Match 15.2%; Score 459; DB 2; Length 491;
Best Local Similarity 26.6%; Pred. No. 6.8e-28;
Matches 122; Conservative 64; Mismatches 130; Indels 142; Gaps 12;
QY 182 QYFQKLGRCVRSVNTANVTLGQPMLEVTVYRHG--RAYVPIAQVKDVVVTDQIPV 240
Db 1 QWQVQLGPGVSLGSLGTDKAMLGTYNMEVTVYHRRGSRVYVPLAHSSAFTITDQVPSV 60
QY 241 TMFQKDRNSDSEFLKDLPIPMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 61 SVSQALDGRNKRFLRKQPLTFALQLHDPSPGLYAGADLSYTWDFGDSGTGLISRALTY 120
QY 301 HTYVLNGTFSLNLTVKAAP-----GPCPPPPPPPP----- 331
Db 121 HTYLSGPPVTAQVVLQAAIPLTSCGSPVPFGTTDRHVTAEAPGTAGQVPTTEVMTTP 180
QY 332 -----PS 333
Db 181 GOVPTAEAPGTGVGVPTTEDVGTTPQVATSKVLSTTPEVMTAKAGRTPEVSTEPS 240
QY 334 KPTPSLG-----PAGDN-----PLE-----LSRIPDE----- 355
Db 241 GTTVTQGTTELPEVTTAGEVSTPEAGSNYSFMTEGTAGSLSPDPDDTATLVLEKRA 300
QY 356 ---NCQINRYGHFOATITIVGILEVNIQMTDVLMPVPWPBESSLIDFVVTCCGSIPTVC 413
Db 301 PLDCVLYRYSFSLTDIVS-----IESAEILOQVSSSGDEGAFELTVSCQGLPKAE 353
QY 414 TIISDPTCEITQNTVCSVDVDEMCLLTVRTE-NGSGTYCVNLTGDDTSLALTSLIS 472
Db 354 MDISPGCOLPAQLCQVPPSPACQLVLHQLKGGSGTYCLNVSLADTNSLAVSTQ 413
QY 473 VPRDRPASPLRMANSALISVGCCLAIFVT--VISLVYKK 509
Db 414 MFGQ--AGLRQAP---LFGVIGILLVLTALLASLIYR 446

RESULT 6

A41105
protein-tyrosine-phosphatase (EC 3.1.3.48) PTPN4, nonreceptor type 4 [validated] - human
N/Alternate names: PTPase MEG
C/Species: Homo sapiens (man)
C/Date: 20-Mar-1992 #sequence_revision 02-May-1994 #text_change 05-Oct-2004
C/Accession: A41105
R/Gu, M.; York, J.D.; Warshawsky, I.; Majerus, P.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 5867-5871, 1991
A/Title: Identification, cloning, and expression of a cytosolic megakaryocyte protein-ty
A/Reference number: A41105; MUID:91288564; PMID:1648233
A/Molecule type: mRNA
A/Accession: A41105
A/Residues: 1-926 <GUA>
A/Cross-references: UNIPROT:P29074; UNIPARC:UPI0000000CA9; GB:M68941; NID:g190747; PIDN:
A/Experimental source: megakaryocytes, cell line MEG-10
C/Genetics:

A/Gene: GDB:PTPN4
A/Cross-references: GDB:131387; OMIM:176878
A/Map position: 9q31-9q31
C/Superfamily: protein-tyrosine phosphatase, non-receptor types 3/4; GLGF domain homolog
C/Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F/31-308/Domain: protein 4.1 membrane-binding domain homology <B41>
F/523-597/Domain: GLGF domain homology <GLG>
F/679-900/Domain: protein-tyrosine-phosphatase homology <PTP>
F/852/Active site: Cys (phosphocysteine intermediate) #status predicted
F/858/Binding site: substrate phosphate (Arg) #status predicted

Query Match 4.5%; Score 136; DB 1; Length 926;
Best Local Similarity 22.0%; Pred. No. 0.025;
Matches 87; Conservative 43; Mismatches 125; Indels 140; Gaps 17;

Qy 40 EHNOLN-----GWSSDENWNEKLYPWKGMWKNWKGRRVQAVLTSDSP 87
Db 204 EFNLYNTARTLELVGFYARDQSN-NEIMIGVMSGELLYKN-----RVRMNTFP 254
Qy 88 ALVGSNITFAVNLIFPRCKEDANG-----NTVYEKNCRNEAGLSADPVYVNTWATSE 140
Db 255 WLKIVKISFKCKQFFIQLRKELHESRETLGLFNVMVYRACKN-----LWKACVE 303
Qy 141 DSDGENTGQSHNVFPDGPFPHPHGWRRWNFIYVHTLGYQFKLGRCSV----- 192
Db 304 -----HHTFFRLDRPLPQK-----NFFAHYFTLGSKFRYCGRTVEQSVQYK 347
Qy 193 -----RVSVNTANTVLGPQLMEVTVRRHGRAYVPIAQVKDVVYVTDQIPVFVTMFQKND 247
Db 348 KANKDRVFARSPSKPLARKLMDWEVSRN-----SISDRLETQSLRSRSPGTFNH 399
Qy 248 RNSS---DETLKDLPI--MFDVLIIH-DPSH-FLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 400 RNSFTQTGEGTLRSPSSVGHLDVHMVHTSPSEVFVNQR-----SPSSTQA 443
Qy 301 HTYVLNGTFSNLNTVKAAPGCPGPPPPPPR----- 331
Db 444 NSIVLESSPS-----QETPDGKPPALPPKQSKNSWNQIYHSHSQDLESINETFDI 497
Qy 332 ---PSKTPSLGPGAGDNPLELSRIPDENCQINRYG 363
Db 498 PSSKEKPTPGGIPHDNLVLIRMKPDEN---GRFG 529

RESULT 7

T07907
hydroxyproline-rich glycoprotein GAS28 precursor - Chlamydomonas reinhardtii
C/Species: Chlamydomonas reinhardtii
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C/Accession: T07907
R/Rodriguez-Martinez, H.; Haring, M.A.; Von Gromof, E.; Beck, C.F.
submitted to the EMBL Data Library, July 1997
A/Reference number: Z16207
A/Accession: T07907
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A/Residues: 1-446 <ROD>

A/Cross-references: UNIPROT:O22458; UNIPARC:UPI000000A8A0P; EMBL:AF015883; NID:g2384727;
A/Experimental source: gametes
C/Genetics:

A/Gene: GAS28

F/1-29/Domain: signal sequence #status predicted <SIG>

F/30-446/Product: hydroxyproline-rich glycoprotein GAS28 #status predicted <MAT>

Query Match 4.0%; Score 120.5; DB 2; Length 446;

Best Local Similarity 29.5%; Pred. No. 0.15; 38; Indels 46; Gaps 6;

Matches 38; Conservative 7; Mismatches 7; Mismatches 46; Gaps 6;

Qy 320 PGFCPPPPPPPPRPSKPTPSLGPA-----GDNFLELSRIPDENCQINRYGHFQ 366

Db 240 PPSPPPPPPPTSPPELPPAPQAPARKRPPPPASPPPSRSDPFCQQRNARGSRL 299

Qy 367 AT-----ITIVGILEVNI-IQMTDVLMPFVWPFESSLIDFVTCQGSIPTEVC----- 413

Db 300 MTTASNNITVVGILTRICFNVALKD-----CEN--PNSKCCFEFLYK 340

Qy 414 -TIISDPTC 421

Db 341 MEIETDPTC 349

RESULT 8

C64483

hypothetical protein MJ1468 - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C/Accession: C64483

R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

rsen, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: C64483

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1009 <BUL>

A/Cross-references: UNIPROT:Q58863; UNIPARC:UPI000013AB23; GB:U67587; GB:L77117; NID:g15

C/Genetics:

A/Map position: FOR1437031-1440060

Query Match 3.8%; Score 114.5; DB 2; Length 1009;

Best Local Similarity 18.3%; Pred. No. 1.4;

Matches 88; Conservative 58; Mismatches 157; Indels 177; Gaps 19;

Qy 100 LIIPRCQKEDANGNIVYEKNCRNEAGL-----SADPVVY----- 133

Db 490 LYPHVFIIDNGNI-----EVGIPPIGVGGYSSYQIYASPIATPIYPIINITIV 541

Qy 134 ---NWT-AWSESDSGENGTCQSHNVFPDGPFPHPHGWRRWNFIYVHTLGYQFKLGR 189

Db 542 EPASWTWYV-----HHIYFGDGSVWKPKKGPYTFHTVTSBGVY----- 582

Qy 190 CSVRVSVNTANTVLGPQLMEVTVRRHGRAYVPIAQVKDVYVTDQIPVFVTMFQKNDRN 249

Db 583 -PIYMKVYTAE-----NMKTVTVVIDNKNPIAKLYTYPNPAS 617

Qy 250 SSDE-----TFKDLPIM-----FDVLHDPSPHFLN 275

Db 618 YKDTISPSINSDPDANRMIPENYVYGTILGYIPISPNSPMARIYGFNLTYVD----- 671

Qy 276 YSTINYKWSFGDNTGLFVSTNHTV-NHT---YVLNGTFSNLNTVKAAPGCPPPP----- 326

Db 672 -SNGNVAWNYSSNELTIISKSPFGNVTAKLVWDGNGVNSTTVKFSVINRPPVQAQFIY 730

Qy 327 -PPPREFSKP-----TPSLGPGAGDNPLELSRIPDENC-----QINRVGHFQATTTI 371

Db 731 YPDKPEPNEDVRFVSQSYDPEGEIAYIWNFGDGTIVNTTDTIVHHKYERPGYVTVTLTV 790

Query Match 3.6%; Score 109.5; DB 2; Length 348;
Best Local Similarity 21.2%; Pred. No. 0.76;
Matches 83; Conservative 51; Mismatches 128; Indels 129; Gaps 22;
QY 61 VKKGDWRKNSWG--GRVQAVLTSQSPALVGNITFAVNLIPRCKEDANGNIVYEK 118
D 16 VFETGQTKKLSATTGNDGTATAIIRN--AAGGGTIVTVCVL-----DADGKV--SD 63
QY 119 NCR-NEAGLSADPVYVNTWAKESDGEENGQSHHNVDPGKPPHHGWRNRNFIYVF 177
D 64 QCDVHFAVASAD--IISFSVSGATDPLGYSNNKGA-ARVFAEIIIPSPNRP---VTNCLINF 117
QY 178 HTLG-----QYFQKLGRCVRSVNTANVTGLPQLMVEVTVRRHGRAYVPIAQVKDYYV 231
D 118 STFGVSGIINNYVNIIGEPN-----SVSSGSLMQ-----FVQIGNYKTT-- 157
QY 232 VTDQIPVVTWFOKNDNRNSDETFLKDLPIVDFVLIHDPHFNLNSTINYKWSFGDNTGL 291
D 158 ----IPV-----DVSYKGPE-----EGSVGIQAAL 178
QY 292 FVSTNHTVNTVYVLTNGFSL---NLTVKAAAGPCPPPPPPRPSKPTSLGPGADNPLE 348
D 179 VTTTQP-----DKAFALSTPNARVDFINPAP-PPPPPPSPCPPPPPPPPPPPFVR 229
QY 349 LS-----RIPDENCQINRYGHFQATITIVEGILEVNIQMTDVLMPVPWPBESS 396
D 230 PSYTIQIDAVVDNIPITPGRGDIARVTVFQNGRPLAGALVQCSLVSLSG-----NTS 281
QY 397 LIDFVVTCCGSIPEVCTTIISDPCEITQNT 427
D 282 LM-FALAA-----NEIANLVSP-----QNT 301
RESULT 12
Ti7414
TipC protein - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T17414
R;Steger, J.T.; Laub, M.T.; Loomis, W.F.
submitted to the EMBL Data Library, July 1998
A;Description: Interaction of tip genes in early Dictyostelium discoideum development.
A;Reference number: 218774
A;Accession: T17414
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3848 <STE>
A;Cross-references: UNIPROT:O76737; UNIPARC:UPI0000076F0B; EMBL:AF079445; NID:G3420744;
A;Experimental source: strain AX4
C;Genetics:
A;Gene: tipC
A;Introns: 72/3
C;Function:
A;Description: required for tip formation
Query Match 3.6%; Score 108; DB 2; Length 3848;
Best Local Similarity 19.0%; Pred. No. 28;
Matches 69; Conservative 55; Mismatches 132; Indels 108; Gaps 14;
QY 3 CLYYFLGFLLAARLPDAKRFDHVLGNERPSAYMREHNLQNGSSDENDWNEKLYPVW 62
D 918 QQYVNLVNIQSVTPIDQQQ---LIGSE--SMYSAFDYSVYQSNIFKQPQOLQPOQ 972
QY 63 KRGMWRKNSWKGROVALTSDSPALVGSNITFAVNLIPRCKEDANGNIVYEKCRN 122
D 973 KQ-----QSPPLTSSP-----PLIKKKKSYSKFIDSYKMF 1003
QY 123 EAGLSADPY-----VYNTWATWESDSDGENTGQSH-----HNVFDPGKPPHHPGWR 170
D 1004 EASNIDKFNHLRLDNYDSSGSSNNDSSNTTTNNQLVSHNV----- 1048
QY 171 WNFIVFHTLGGYFQKLGRCVRSVSN---TANVTGLPQLMVEVTVRRHGRAYVPIAQVK 227

Db 1049 -ALVYLRCLHGTYVQKLFDTNLELLIKGMWIEDCFQKPSI-----AHGSG----- 1093
QY 228 DVYVVTQDIPFVFTMQCKDRN-----SSDETFKDLPIPMEDVLHDP 270
D 1094 -----GDLATTTNRFPKPDVNTFGGSGNENLIFKIQIQQSSDSPFYNNIDKLLDIELSQI 1148
QY 271 SHFLANTSTINYKWSFGDN-TGLFVSTNHTVNTVTLNGTFSNLITVKAAPG---PCPPP 326
D 1149 NLILNRKTVAGLIEFSSVSNLSILKXNNQNNQ---NNNQNNQNNINESFTVITSPPP 1205
QY 327 PPPP 330
D 1206 PPPP 1209
RESULT 13
T23694
hypothetical protein M03C11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23694
R;McMurray, A.
submitted to the EMBL Data Library, April 1995
A;Reference number: Z19783
A;Accession: T23694
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-848 <WIL>
A;Cross-references: UNIPROT:Q21489; UNIPARC:UPI0000079560; EMBL:Z49128; PIDN:CAA88959.1;
A;Experimental source: clone M03C11
C;Genetics:
A;Gene: CESP-M03C11.2
A;Map position: 3
A;Introns: 113/2; 147/3; 185/3; 379/1; 482/3; 553/2; 688/3; 762/3
Query Match 3.5%; Score 106.5; DB 2; Length 848;
Best Local Similarity 20.8%; Pred. No. 4.5;
Matches 117; Conservative 75; Mismatches 203; Indels 167; Gaps 30;
QY 83 TSDSPALV--GSNITFAVN-----LIFFRCKEDANGNIVYEKCRNEAGLSADP 130
D 206 TRFQPRIVTCASTRGLTCVNEEVKKLKLNHLINELKMLKNGMSEKEKVQKLEKGTTKT 265
QY 131 YV-----YNTWATWESDSDG-----ENGTGOSHNVFPDGKPF----- 162
D 266 KTCATSCFYNSTQIEDVNVNGVLSNKLSTLEVSQKGLNSGCPYFATRKSVPCQQLVLL 325
QY 163 PH-----HPGWRR-W-----NFTY-----VFHTLQYF-----QKLGK-----CSV-RVS 195
D 326 PYQVLLHDGTRKAWGIELKDNVIVLDEAHNVLTISFFFRKLEKYSKNGFKNCISYAE 385
QY 196 VNTANVTGLPQLMVEVTVRRHGRAYV---PIAQVKDYYVVTQDIPFVFTMQCKDRN 252
D 386 ISTKSLTLALRL--IREYNAHYKLLAHLNLLYMKQLESLSKMLIFLNSQSKEDVNTMA 443
QY 253 ETEFLKDLPIVDFVLIHDPHFNLNSTINYKWSFGDNTGLFVSTN-----HTVNTVYVLTNGT 308
D 444 Q-LARNLNLIL-EINLFLKAEYMEKTLDCCKFH-----GFYMLQKEIKKENEKPKLTGI 496
QY 309 FSNLNTVKAAPGCPPPPPPPRPSKPTP-----SLGPAGDNPLELSRIPDENQINRYG 363
D 497 QKL-MAAKEAPEPEAEPLPPKP-VPSPLFSLSKSFIDALTNKCEDGRILIVEKSATKPF 554
QY 364 HF-----QATITIVEGILEVNIQMTDVLMPVPWPBESSLIDFVVTCCGS 407
D 555 RFMLLNPADRLSEVWTSARATI-LVGGTME-----PAQLLVE--TLSRGS 596
QY 408 IPEVCTTIISDPCEITQNTVCSFVDVDEMCLLTGRTENGSGTYCVNLTLGDDTSALT 467
D 597 IGAD-----SIRRFSCCHVIDDQLAVTVERTVDGKPFQLTQYTRGADTTLURL 646
QY 468 STLISVPDRDPASPRLMANSALIS--VGCLAI FVTWISLL--VYKHKHENPIENSPGNV 523

```

Db      647  ATSI-----QALPHIPNGWVFPSYDFLNFQKKMEF----- 681
          :||      ||||      :|||      |      ||||
Qy      524  VRSKGLSVFLNRAKAVFPFGNQ 545
          |      :      ||||
Db      682  ----GILKRIIEKIAVFTESRQ 699
          :      :      ||||

```

RESULT 14

QJ0533
Genome polyprotein - Kennedyya yellow mosaic virus (strain Jervis Bay)
N/Alternate names: RNA nucleotidyltransferase [RNA-directed]; RNA replicase
N/Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C/Species: Kennedyya yellow mosaic virus
C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C/Accession: JQ0533
R/Ding, S.; Keese, P.; Gibbs, A.
J. Gen. Virol. 71, 925-931, 1990
A/Title: The nucleotide sequence of the genomic RNA of Kennedyya yellow mosaic tymovirus
A/Reference number: JQ0532; MUID:90218040; PMID:2324710
A/Accession: JQ0533
A/Molecule type: genomic RNA
A/Residues: 1-1874 <DIN>
A/Cross-references: UNIPROT:P36304; UNIPARC:UPI0000131EA7; GB:D00637; NID:G221969; PIDN:
C/Superfamily: eggplant mosaic virus RNA-directed RNA polymerase
C/Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; R
F/1002-1009/Region: nucleotide-binding motif A (P-loop)
F/1064-1069/Region: nucleotide-binding motif B
F/1008/Binding site: ATP (Lys) #status predicted

Query Match 3.5%; Score 106; DB 1; Length 1874;
Best Local Similarity 20.2%; Pred. No. 15;
Matches 75; Conservative 38; Mismatches 100; Indels 158; Gaps 16;

Qy 257 KQLPIMFDVLIHDPSPHFLNYSTIN-----YKWSFGDNTGLFVS 294
|||:::
Db 526 KQIPLL-----PR--MSEFTVNLPLQPPMMLAIGASLVPELAFLLSWLSGD-----VDLO 573
|||:::

Qy	295	TNHTVNHVTYVLNGTFFSLNL.T--VKAAAPGP-----CP-PPPPPPRPS	333
Db	574	TCCHDIYHHHLHPENFTLSWTRTPYLALASPFLPYAHSPPLPPLPVNSSPLPFPFPPPPPPPS	633

Qy 334 KPTPSLGP- AGDNPLELSRIPE 355
:
Dd 634 OPPLSGPATAPSAOPTGEPILLAPTTELKPSSNNPNPNPSSAGNPPKSSSDN 693

Qy 356 NCQINRYGHFOATIT-----IVEGILEVNII-----QMTDVLMPVP 391
| : | : | : | :
Db 694 PRAPNKPTPTSSSTTPPSNII.PLOGSTHSPFT.SDGQI.NVYSAT.DDDDDNNNTTTSII.LPDP 753

Qy 392 WPESLLIDFVWTCQGSIPTEV-CTIIISDPTCEITQNTVCSPPVDDEMCLLTIVRRTFNGSG 450

Dh 754 KP-----PPEVCSGLMADPTC-----VCDAYSCSS--LVDPPEPNTA 789

QY 451 TVCNVLTIGDDTSIALTST--LISVPDRDPASPLR-----MANSALISVGC 494

Dh : : : : : :

790 SPTTLPETSPDTPDMKONCTATAYAGSRWINDHPIWHNCEIQTWPDRIISNCPSIDSCVM 840

RESULT 15

Protein F21D18.18 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: C96521
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Anderson, N.F.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-492 <STO>
A;Cross-references: UNIPROT:Q9LNG3; UNIPARC:UPT00000AA2C0; GB:AE005173; NID:g8778527; P: C;Genetics:
A;Gene: F21D18.18
A;Map position: 1

Query Match 3.5%; Score 105; DB 2; Length 492;
Best Local Similarity 24.7%; Pred. No. 2.8;
Matches 41; Conservative 10; Mismatches 43; Indels 72; Gaps 8;

Qy	272	HFLNYS	TIN	YKWS	FGDNTGLFVSTNH	TVNHTVVLNGTFSLNLTKVAAPGCP	PPPPPPPR	331
		:						
		:						
Db	20	HFTNVAR	ARY	YCHKG	-----RGVTH-----	-----PUPPPPPPPL		50
		:						
		:						

Qy	332	-----	PKKPTPSLGPA-GDNP	-----	LELSRIPDENCQ	-----	358
Db	51	ETANPPDOVP	SDYPSAPDPA	PGDSISGCFV	VTSEGAWGDCS	DDTAFAFODAWKAACAVE	110

Search completed: December 24, 2005, 04:36:31
Job time : 48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:25:06 ; Search time 233 Seconds
(without alignments)
1695.690 Million cell updates/sec

Title: US-09-853-880A-17
Perfect score: 3026
Sequence: 1 MECLYYFLGILLARLPD.....FPGNQEKDPLLNQEFKGV 560

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	3010	99.5	572 1 GPNMB_HUMAN	Q14956 homo sapien
2	2197	72.6	574 1 GPNMB_MOUSE	Q99p91 mus musculus
3	2179	72.0	572 1 GPNMB_RAT	Q6p7c7 rattus norv
4	2163	71.5	572 2 Q9QZF6_RAT	Q9qzf6 rattus norv
5	2042	67.5	526 2 OB8VA0_MOUSE	Q8bva0 mus musculus
6	1588.5	52.5	559 1 QNR71_COTJA	Q90372 coturnix co
7	1431.5	47.3	595 2 Q4V805_XENLA	Q4v805 xenopus lae
8	1029	34.0	206 2 Q96F58_HUMAN	Q96f58 homo sapien
9	1022	33.8	206 2 Q8IXJ5_HUMAN	Q8ixj5 homo sapien
10	843	27.9	538 2 Q4RKH9_TETNG	Q4rkh9 tetraodon n
11	738.5	24.4	719 2 Q4KLIV9_XENLA	Q4kliv9 xenopus lae
12	736.5	24.3	721 2 Q6DDN6_XENLA	Q6ddn6 xenopus lae
13	729.5	24.1	746 2 Q6DIR2_XENTR	Q6dir2 xenopus tro
14	682.5	22.6	808 2 Q52KZ4_XENLA	Q52kz4 xenopus lae
15	671.5	22.2	626 2 Q9CZB2_MOUSE	Q9czb2 mus musculus
16	667.5	22.1	845 2 Q4VW61_BRARE	Q4vw61 brachydanio
17	663.5	21.9	636 2 Q4LE84_PIG	Q4le84 sus scrofa
18	659	21.8	626 1 PHE17_MOUSE	Q60696 mus musculus
19	652	21.5	604 2 Q4RIM0_TETNG	Q4rim0 tetraodon n
20	651	21.5	733 2 Q93391_COTCO	Q93391 coturnix co
21	650	21.5	763 1 PHE17_CHICK	Q98917 gallus gall
22	649.5	21.5	703 2 Q4R7T2_MACFA	Q4r7t2 macaca fasc
23	622	20.6	661 1 PHE17_HUMAN	Q40967 homo sapien
24	561	18.5	413 2 Q4S693_TETNG	Q4s699 tetraodon n
25	459	15.2	491 1 PHE17_BOVIN	Q06154 bos taurus
26	442.5	14.6	251 2 Q5R2V2_BRARE	Q5r2v2 brachydanio
27	399	13.2	461 2 Q97884_HORSE	Q97884 equus cabal
28	247.5	8.2	236 2 Q9QY67_MOUSE	Q9qy67 mus musculus
29	182.5	6.0	423 2 Q5R6F5_PONPY	Q5r6f5 pongo pygma
30	182	6.0	397 2 Q8N3R2_HUMAN	Q8n3r2 homo sapien
31	181.5	6.0	423 2 Q8N0W9_HUMAN	Q8n0w9 homo sapien

32	181	6.0	435	2	Q8N3G9_HUMAN	Q8n3g9 homo sapien
33	167	5.5	141	2	Q9QY70_MOUSE	Q9qy70 mus musculus
34	156	5.2	419	2	Q6NXM3_MOUSE	Q6nxm3 mus musculus
35	155.5	5.1	3779	2	Q4RZY1_TETNG	Q4rzy1 tetraodon n
36	153	5.1	321	2	Q8IY46_HUMAN	Q8iy46 homo sapien
37	144	4.8	1988	2	Q8TP21_METAC	Q8tpz1 methanosarc
38	142.5	4.7	1817	2	Q8TI59_METAC	Q8ti59 methanosarc
39	140	4.6	688	2	Q8TR88_METAC	Q8tr88 methanosarc
40	137	4.5	906	2	Q8TPV9_METAC	Q8tpv9 methanosarc
41	136	4.5	926	1	PTN4_HUMAN	P29074 homo sapien
42	133	4.4	881	2	Q6H7U3_ORYSA	Q6h7u3 oryza sativ
43	132	4.4	331	2	Q4V9B2_BRARE	Q4v9b2 brachydanio
44	127.5	4.2	603	2	Q7U5X8_SYNFX	Q7u5x8 synecococc
45	126.5	4.2	514	2	Q7PYC8_ANOGA	Q7pyc8 anopheles g

ALIGNMENTS

RESULT 1
GPNMB_HUMAN
ID GPNMB_HUMAN STANDARD; PRT; 572 AA.
AC Q14956; Q6UVX1; Q8N1A1;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane glycoprotein NMB precursor (Transmembrane glycoprotein HGFIN).
GN Name=GPNMB; Synonyms=HGFIN, NMB; ORFNames=UNQ1725/PRO9925;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
RC TISSUE=Melanoma;
RX MEDLINE=95113576; PubMed=7814155;
RA Waterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J.,
RA van Muljen G.N.P., Ruiter D.J., Bloemers H.P.J.;
RT "NMB, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts.";
RL Int. J. Cancer 60:73-81 (1995).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).
RC TISSUE=Peripheral blood;
RX MEDLINE=22498106; PubMed=12609765; DOI=10.1016/S0167-0115(02)00288-4;
RA Bandari P.S., Qian J., Yehia G., Joshi D.D., Maloof P.B., Potian J.,
RA Oh H.S., Gascon F., Harrison J.S., Rameshwar P.;
RT "Hematopoietic growth factor inducible neurokinin-1 type: a transmembrane protein that is similar to neurokinin 1 interacts with substance P.";
RL Regul. Pept. 111:169-178 (2003).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klinkowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seehagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";
RL Genome Res. 13:2265-2270 (2003).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,

RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
 RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
 RA Fowell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
 RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
 RA Ozerly P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
 RA Kozlowsky-Reilly A., Leonard S., Rohlffing T., Rock S.M.,
 RA Tin-Wollam A.-M., Abbott A., Minx P., Maupin R., Strommatt C.,
 RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
 RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.B.,
 RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
 RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
 RA Gillett W., Zhou Y., James R., Phelps K., Iadamoto S., Bubb K.,
 RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Purrey T.S.,
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
 RA Waterston R.H., Wilson R.K.;
 RT "The DNA sequence of human chromosome 7."; ;
 RL Nature 424:157-164 (2003). ;
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RC TISSUE=Brain;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X.J., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q14956-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q14956-2; Sequence=VSP_013001;
 CC -!- TISSUE SPECIFICITY: Not restricted to the melanocytic lineage.
 CC -!- DEVELOPMENTAL STAGE: Expression in poorly metastatic melanoma cell
 CC lines; no expression in highly metastatic melanoma cell lines.
 CC -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
 CC -!- SIMILARITY: Contains 1 PKD domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; X76534; CRA54044.1; -; mRNA.
 DR EMBL; AF322909; AAG42839.1; -; mRNA.
 DR EMBL; AY359124; AAO89481.1; -; mRNA.
 DR EMBL; AC005082; AAP22336.1; -; Genomic_DNA.
 DR EMBL; BC032783; AAH32783.1; -; mRNA.
 DR PIR; I38065; I38065.
 DR Ensembl; ENSG00000136235; Homo sapiens.

DR HGNC; HGNC:4462; GPNMB.
 DR MIM; 604368; -.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
 DR InterPro; IPR000601; PKD.
 DR SMART; SM00089; PKD; 1.
 DR PROSITE; PS00093; PKD; 1.
 KW Alternative splicing; Glycoprotein; Polymorphism; Signal;
 KW Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 572 Transmembrane glycoprotein NMB.
 FT TOPO_DOM 22 486 Extracellular (Potential).
 FT TRANSMEM 487 507 Potential.
 FT TOPO_DOM 508 572 Cytoplasmic (Potential).
 FT DOMAIN 240 327 PKD.
 FT MOTIF 64 66 Cell attachment site (Potential).
 FT COMPTAS 320 332 Poly-Pro. (Potential).
 FT CARBOHYD 93 93 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 146 146 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 200 200 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 249 249 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 275 275 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 300 300 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 306 306 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 312 312 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 467 467 N-linked (GlcNAc...) (Potential).
 FT VARSPPLIC 340 351 Missing (in isoform 2).
 FT /FTID=VSP_013001.
 FT VARIANT 195 195 S -> C (in dbSNP:530436).
 FT /FTID=VAR_012076.
 FT VARIANT 197 197 N -> H (in dbSNP:530413).
 FT /FTID=VAR_012077.
 FT CONFLICT 354 354 A -> T (in Ref. 3).
 FT SQ SEQUENCE 572 AA; 63923 MW; 2465C12CF0F3996 CRC64;
 Query Match 99.5%; Score 3010; DB 1; Length 572;
 Best Local Similarity 97.9%; Pred. No. 1.le-218;
 Matches 560; Conservative 0; Mismatches 0; Indels 12; Gaps 1;
 Qy 1 MECLYPLGLFLLAARLPLDAARFHDVGLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
 Db 1 MECLYPLGLFLLAARLPLDAARFHDVGLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
 Qy 61 VVKRGDMRWKNSKGGVQAVLTSDSPALVGSNTTFAVNLIFPRCKEDANGNIYKNC 120
 Db 61 VVKRGDMRWKNSKGGVQAVLTSDSPALVGSNTTFAVNLIFPRCKEDANGNIYKNC 120
 Qy 121 RNBAGLSADPYVYNWNTAWSDESDGNGTQSHNVFPDGPFPHPHGWRRWNFIYVFTLL 180
 Db 121 RNBAGLSADPYVYNWNTAWSDESDGNGTQSHNVFPDGPFPHPHGWRRWNFIYVFTLL 180
 Qy 181 GQYFQKLGRCSRVSVNTANTVIGPOLMEVTVYRRHGRVYVIAQVXDYVVTQIPVFV 240
 Db 181 GQYFQKLGRCSRVSVNTANTVIGPOLMEVTVYRRHGRVYVIAQVXDYVVTQIPVFV 240
 Qy 241 TMQKNDNSDETFLLKDLPIMEDVLIDHPSHFLNYSTINYKSPGNTGLFVSTNHTVN 300
 Db 241 TMQKNDNSDETFLLKDLPIMEDVLIDHPSHFLNYSTINYKSPGNTGLFVSTNHTVN 300
 Qy 301 HTYVLNGTFFSLNLTAKAAAGPCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 348
 Db 301 HTYVLNGTFFSLNLTAKAAAGPCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360
 Qy 349 LSRIPDENCQINRYGHFQAATITVEGILEVLENIQMTDVLMPVPBPSSLDIFVTCQSI 408
 Db 361 LSRIPDENCQINRYGHFQAATITVEGILEVLENIQMTDVLMPVPBPSSLDIFVTCQSI 420
 Qy 409 PTEVCTIIISDPTCEITONTVCSPVDVDEMCLLTVRRRTFNGSGTYCVNLTLGDDTSLA 468
 Db 421 PTEVCTIIISDPTCEITONTVCSPVDVDEMCLLTVRRRTFNGSGTYCVNLTLGDDTSLA 480

QY 469 TLISVDPDRPASPPIRMANSLISVCLAFVTVISLLVYKKHKNYPIENSPGNVVRSGK 528
 DB 481 TLISVDPDRPASPPIRMANSLISVCLAFVTVISLLVYKKHKNYPIENSPGNVVRSGK 540

QY 529 LSVFLNRAKAVFFPGNQEKDPLLNQKQEPKGV 560
 DB 541 LSVFLNRAKAVFFPGNQEKDPLLNQKQEPKGV 572

RESULT 2

GNMB MOUSE
 ID GPNMB MOUSE STANDARD; PRT; 574 AA.
 AC Q999J1; Q8BV9; Q8X14; Q9QXA0;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Transmembrane glycoprotein NMB precursor (Dendritic cell-associated
 DE transmembrane protein) (DC-HIL)
 GN Name=Gpmb; Synonyms=Dchil, Nmb;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BA1B/c;
 RX MEDLINE=21269372; PubMed=11114299; DOI=10.1074/jbc.M008539200;
 RA Shikano S., Bonkobara M., Zukas P.K., Arizumi K.;
 RT "Molecular cloning of a dendritic cell-associated transmembrane
 RT protein, DC-HIL, that promotes RGD-dependent adhesion of endothelial
 RT cells through recognition of heparan sulfate proteoglycans.";
 RL J. Biol. Chem. 276:8125-8134 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
 RX PubMed=12638126; DOI=10.1016/S1567-133X(02)00012-1;
 RA Bachner D., Schroder D., Gross G.;
 RT "mRNA expression of the murine glycoprotein (transmembrane) nmb
 RT (Gpmb) gene is linked to the developing retinal pigment epithelium
 RT and iris.";
 RL Gene Expr. Patterns 1:159-165 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Retina, and Skin;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gusninch S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kogut A.A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKensie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Whitehead C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Alzawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ichi Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=FVB/N; TISSUE=Salivary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: Could be a melanogenic enzyme (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Vesicular,
 CC endosomal like structures.
 CC -1- TISSUE SPECIFICITY: In early development mRNA is detected at high
 CC levels in the outer layer of the retina. Later in development and
 CC expression gets restricted to the retinal pigment epithelium and
 CC iris.
 CC -1- SIMILARITY: Belongs to the Pmel-17/NMB family.
 CC -1- SIMILARITY: Contains 1 PKD domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; AF322054; AAK14240.1; -; mRNA.
 CC EMBL; AJ251685; CAB65272.1; -; mRNA.
 CC EMBL; AK044764; BAC32074.1; -; mRNA.
 CC EMBL; AK076347; BAC36306.1; -; mRNA.
 CC EMBL; BC026375; AAH26375.1; -; mRNA.
 CC Ensembl; ENSMUSG0000029816; Mus musculus.
 CC MGI; MGI:1934765; Gpmb.
 CC GO; GO:0005615; C:extracellular space; TAS.
 CC GO; GO:0005887; C:integral to plasma membrane; IDA.
 CC GO; GO:0008201; F:heparin binding; IDA.
 CC GO; GO:0005178; F:integrin binding; IDA.
 CC GO; GO:0007155; P:cell adhesion; IDA.
 CC InterPro; IPR000601; PKD.
 CC SMART; SM00089; PKD; 1.
 CC PROSITE; PS50093; PKD; 1.
 KW Glycoprotein; Signal; Transmembrane.
 FT SIGNAL 1 22 Potential.
 FT CHAIN 23 574 Transmembrane glycoprotein NMB.
 FT TOFO DOM 23 502 Extracellular (Potential).
 FT TRANSMEM 503 523 Potential.
 FT TOPO DOM 524 574 Cytoplasmic (Potential).
 FT DOMAIN 250 338 PKD.
 FT MOTIF 558 560 Cell attachment site (Potential).
 FT CARBOHYD 93 93 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 200 200 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 249 249 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 275 275 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 300 300 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 306 306 N-linked (GlcNAc...) (Potential).

```
FT CARBOHYD 312 312 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 463 463 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 471 471 N-linked (GlcNAc... ) (Potential).
FT CONFLICT 36 36 N -> D (in Ref. 2 and 3; BAC36306).
FT CONFLICT 333 333 SPP -> S (in Ref. 2).
FT CONFLICT 339 341 SPP -> LPS (in Ref. 2).
FT CONFLICT 498 498 L -> M (in Ref. 3; BAC32074).
FT CONFLICT 541 541 K -> E (in Ref. 3; BAC32074).
SQ SEQUENCE 574 AA; 63675 MW; 6BAE9AC27AE6ACD0 CRC64;

Query Match 72.6%; Score 2197; DB 1; Length 574;
Best Local Similarity 70.7%; Pred. No. 2.9e-157;
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

Qy 1 MECIYFLGLLLAARLPDAARFHDVILGNRPSAYMRHNLNGWSSDENWNEKLYP 60
Db 1 MESLCGVLLAAGLPQAARFHDVILGNRPSAYMRHNLNGWSSDENWNEKLYP 60

Qy 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNTWTAWSEDSGNGTQSHHNVDPGKPPHHPGRWNNFYVFTL 180
Db 121 RNEAGLSADPYVYNTWTAWSEDSGNGTQSHHNVDPGKPPHHPGRWNNFYVFTL 180

Qy 121 RNDLGTSDLVYNTWTAGDGDWEDGTSRQHLRFDRRPFPRPHGKMSFYVFTL 180
Db 121 RNDLGTSDLVYNTWTAGDGDWEDGTSRQHLRFDRRPFPRPHGKMSFYVFTL 180

Qy 181 GOYFQKLGRCVRSVNTANTLGPOLMEVTVYRRHGRVYPIAQKDVVVTDOIPVFV 240
Db 181 GOYFQKLGRCVRSVNTANTLGPOLMEVTVYRRHGRVYPIAQKDVVVTDOIPVFV 240

Qy 241 TMFQKNDNRSSDETFLLKDLPIFMDVLIHDPHSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMSQKNDNRSSDETFLLKDLPIFMDVLIHDPHSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 344
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 344

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 344
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 344

Qy 345 NPLEISRTPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 404
Db 345 NPLEISRTPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 404

Qy 361 KSMELSDISNENCRINRYGYFRATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 420
Db 361 KSMELSDISNENCRINRYGYFRATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 420

Qy 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464
Db 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464

Qy 421 KGATPMEACTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 480
Db 421 KGATPMEACTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 480

Qy 465 ALTSTLISVDRDPASPLRMANSALISVGLCIAFVTVISLIVYKHKKEYNPIENSGNVV 524
Db 465 ALTSTLISVDRDPASPLRMANSALISVGLCIAFVTVISLIVYKHKKEYNPIENSGNVV 524

Qy 481 ALTSTLISVDRDPASPLRMANSALISVGLCIAFVTVISLIVYKHKKEYNPIENSGNVV 540
Db 481 ALTSTLISVDRDPASPLRMANSALISVGLCIAFVTVISLIVYKHKKEYNPIENSGNVV 540

Qy 525 RSKGLSVPLNRAKAVFPNGQKDPKLLKNQ 554
Db 525 RSKGLSVPLNRAKAVFPNGQKDPKLLKNQ 554

Qy 541 KRGKGLSVLLSHAKAPFPRGQKDPKLLQDK 570
Db 541 KRGKGLSVLLSHAKAPFPRGQKDPKLLQDK 570

RESULT 3
GNPMB RAT
ID GPNMB RAT STANDARD; PRT; 572 AA.
AC G6P7C7;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transmembrane glycoprotein NMB precursor.
GN Name=Gnmb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
```

```
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BC061725; AAH61725.1; -; mRNA.
DR Ensembl; ENSRNOG0000008816; Rattus norvegicus.
DR RGD; 71008; Gnmb.
DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00093; PKD; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 572 Transmembrane glycoprotein NMB.
FT TOPO_DOM 23 500 Extracellular (Potential).
FT TRANSMEM 501 521 Potential.
FT TOPO_DOM 522 572 Cytoplasmic (Potential).
FT DOMAIN 251 338 PKD.
FT MOTIF 556 558 Cell attachment site (Potential).
FT CARBOHYD 93 93 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 134 134 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 200 200 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 249 249 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 275 275 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 300 300 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 306 306 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 312 312 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 461 461 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 469 469 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 572 AA; 63731 MW; 99854F7773FF946C CRC64;

Query Match 72.0%; Score 2179; DB 1; Length 572;
Best Local Similarity 69.7%; Pred. No. 6.6e-156;
Matches 396; Conservative 74; Mismatches 84; Indels 14; Gaps 1;

Qy 1 MECIYFLGLLLAARLPDAARFHDVILGNRPSAYMRHNLNGWSSDENWNEKLYP 60
Db 1 MESLCGVLLAAGLPQAARFHDVILGNRPSAYMRHNLNGWSSDENWNEKLYP 60

Qy 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNTWTAWSEDSGNGTQSHHNVDPGKPPHHPGRWNNFYVFTL 180
Db 121 RNEAGLSADPYVYNTWTAWSEDSGNGTQSHHNVDPGKPPHHPGRWNNFYVFTL 180

Qy 121 RNDLGTSDLVYNTWTAGDGDWEDGTSRQHLRFDRRPFPRPHGKMSFYVFTL 180
Db 121 RNDLGTSDLVYNTWTAGDGDWEDGTSRQHLRFDRRPFPRPHGKMSFYVFTL 180

Qy 181 GOYFQKLGRCVRSVNTANTLGPOLMEVTVYRRHGRVYPIAQKDVVVTDOIPVFV 240
Db 181 GOYFQKLGRCVRSVNTANTLGPOLMEVTVYRRHGRVYPIAQKDVVVTDOIPVFV 240

Qy 241 TMFQKNDNRSSDETFLLKDLPIFMDVLIHDPHSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRSSDETFLLKDLPIFMDVLIHDPHSHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 346
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 346

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 346
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPAGD 346

Qy 347 LEISRIEDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 406
Db 347 LEISRIEDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 406

Qy 361 MELSDISNENCRINRYGYFRATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 420
Db 361 MELSDISNENCRINRYGYFRATITVEGILEVNIQMTDVLMPVPWESSLIDFVVT 420

Qy 407 SIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 466
Db 407 SIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 466
```

```

Db 421 ATPTACTIISDPTCOIAQRVCSFVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480
Qy 467 TSTLSVPRDRPASPLRMANSALISVGCLAFVTVISLLVYKHKYNPINSPGNVRS 526
Db 481 TSALISIPGKGLSPLRTVNGVLISIGCLAFVTVISLLVYKHKYNPINSPGNVRS 540
Qy 527 KGLSVFLNRAKAVFPFGNQKDPKLNQ 554
Db 541 KGLSVFLSHAKAFPSRGDREKDPKLNQ 568

RESULT 4
OSQZF6 RAT PRELIMINARY; PRT; 572 AA.
AC OSQZF6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Osteoactivin.
GN Name=Gpmb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21611396; PubMed=11746512; DOI=10.1002/jcb.1259;
RA Safadi F.F., Xu J., Smock S.L., Rico M.C., Owen T.A., Popoff S.N.;
RT "Cloning and characterization of osteoactivin, a novel cDNA expressed
in osteoblasts.";
RL J. Cell. Biochem. 84:12-26(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Odgren P.R., Marks S.C. Jr.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Xu J., Safadi F.F., Rosenzweig A.B., Popoff S.N.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184983; AAF03400.1; -; mRNA.
DR Ensembl; ENSRNOG0000008816; Rattus norvegicus.
DR RGD; 71008; Gpmb.
DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS0093; PKD; 1.
SQ SEQUENCE 572 AA; 63772 MW; CB939FB043EC57EB CRC64;

Query Match 71.5%; Score 2163; DB 2; Length 572;
Best Local Similarity 69.4%; Pred. No. 1.1e-154;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

Qy 1 MECLYVFLGLLAAALPLDAKRFHVLGNRPASVYRHNQNGWSSDENNEKLYP 60
Db 1 MESLCGVLFLLAAGLPLQAARFRDVLGHQYDPDRHNNQLRGWSSDENWEQLYP 60
Qy 61 VVKRGDMRWKNSKGGVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYERNC 120
Db 61 VVRGEGRWKDSWEGRVQALTSPPALVGSNITFAVNLIPPRCKEDANGNIYERNC 120
Qy 121 RNEAGLSADPYVYNNTANSESDGNGTQSQSHNVNPPGKPPHPHGRWRNFIYVFTL 180
Db 121 RSDLEASDPYVYNNTGADDEWEDNTSQGHLPDPGKPPRPHGRKKNFYVFTL 180
Qy 181 GQYFQKGRCSVRVSNVNTNLGPMLEVTYVRHGRAYPIAOVKDYVVTDOIPVF 240
Db 181 GQYFQKGRCSVRVSNVNTNLGPMLEVTYVRHGRAYPIAOVKDYVVTDOIPVF 240
Qy 241 TMFQKNDNSDETFLKPLIMFDVLIDHPSHFLNYSINYNKWSFGDNTGLFVSNHTVN 300
Db 241 TMQKNDNSDETFLRDLPIFFDVLIDHPSHFLNYSINYNKWSFGDNTGLFVSNHTLN 300
Qy 301 HTYVLNGTFTSLNLTVKAAAPGCPPPPPPPR-----SKFTPSLGPAGDNP 346

```

```

Db 301 HTYVLNGTFTSLNLTVKAAAPGCPSPSPSSSTSPSPSPSPSPSPSPSPSPSPSP 360
Qy 347 LELSRIEDENCOINRYGHFOATITIVGILRVNIQTMDVLPVWPWPSPSSLIIDFVTCQG 406
Db 361 MELSDISNENCRINRYGVFRATITIVGILRVNIQTMDVLPVWPWPSPSSLIIDFVTCQG 420
Qy 407 SIPTVECTIISDPTCEITONTVCSPVDVDECLITVRRTFNGSGTYCNFTLGGDTSIAL 466
Db 421 ATPTACTIISDPTCOIAQRVCSFVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480
Qy 467 TSTLSVPRDRPASPLRMANSALISVGCLAFVTVISLLVYKHKYNPINSPGNVRS 526
Db 481 TSALISIPGKGLSPLRTVNGVLISIGCLAFVTVISLLVYKHKYNPINSPGNVRS 540
Qy 527 KGLSVFLNRAKAVFPFGNQKDPKLNQ 554
Db 541 KGLSVFLSHAKAFPSRGDREKDPKLNQ 568

RESULT 5
Q8BVA0 MOUSE PRELIMINARY; PRT; 536 AA.
AC Q8BVA0;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mus musculus adult male urinary bladder cDNA, RIKEN full-length
enriched library, clone:9530038P20 product:glycoprotein
(transmembrane) nmb, full insert sequence.
GN Name=Gpmb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schriml L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustacich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RA The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

```


FT CARBOHYD 92 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 133 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 145 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 149 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 192 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 199 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 248 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 274 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 307 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 311 N-linked (GlcNAc...) (Potential)
FT CARBOHYD 459 N-linked (GlcNAc...) (Potential)
SQ SEQUENCE 559 AA; 61714 MW; 8182793BD64A47DD CRC64;
Query Match 52.5%; Score 1588.5; DB 1; Length 559;
Best Local Similarity 53.5%; Pred. No. 2.7e-111;
Matches 297; Conservative 86; Mismatches 157; Indels 15; Gaps 5;
QY 8 LGFLLLAARLPDAAKRFHDLGNRPSAYMRHNLGWSDDNDWNEKLYPVWKRGM 67
DB 8 LALLPAAVLAAMRQDVLNQR-TAPVNNHKKIQGWSDDQWNEKLYPFWEDNDP 66
QY 68 RWKSWKGRVQAVLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLS 127
DB 67 RWKDCWKGGKVTTLVTDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLS 126
QY 128 ADPVYNNWNTANVTLSGQSHNVFPDGPFPHPGWRNFIYVFTLGOYFQKL 187
DB 127 QDQYNNWNTANVTLSGQSHNVFPDGPFPHPGWRNFIYVFTLGOYFQKL 186
QY 188 GRCSVRVSVNTANVTLSGQSHNVFPDGPFPHPGWRNFIYVFTLGOYFQKL 247
DB 187 GRCSVRVSVNTANVTLSGQSHNVFPDGPFPHPGWRNFIYVFTLGOYFQKL 246
QY 248 RNSDTEFLKOLIMFDVLIHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNLG 307
DB 247 RNSDTEFLKOLIMFDVLIHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNLG 306
QY 308 TFSMLTVKAAGPCPPPPPPRPSKTPSL-----GPAGNPLELSRIPDENCOI 359
DB 307 NFTLMLTVQAIIPVPC-KPVTPPSLPTPAVTTDASSNPSAPNAEDN-PDGGCHI 363
QY 360 NRYGHQFQATITVEGILEVNIQMTDVLMPVWPRESSLIDFVTCQSGIPTVCTIISDP 419
DB 364 YRYGYTATITVEGILEVNIQMTDVLMPVWPRESSLIDFVTCQSGIPTVCTIISDP 423
QY 420 TCEITQNTVCSVDVDEMLLTVRTFNGSGTYCVNLTLGDDTSLALSTLISVPDRDPA 479
DB 424 TCQVSGMVCVPVVTDECVLTIRAFDPGTYCINILGDDTSLALSTLISVPDRDPA 480
QY 480 SPLRMANGLISVGCIAIFVTVISLLVYKHKYKNPIENSPGVRSKGLSVFLNRAKAV 539
DB 481 SSSGTTKGVIFLGLLAVGAIGAFVLYKRYKQYKPIERSAGQAEQGLSAYVSNFKAF 540
QY 540 FPGNQEKDPLKQ 554
DB 541 FPKSTERNPLKSK 555

RESULT 7

Q4V805_XENLA PRELIMINARY; PRT; 595 AA.
AC Q4V805;
DT 13-SEP-2005 (Tremblrel. 31, Created)
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCHI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Splice;
RA MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Splice;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Splice;
RX Klein S., Gerhard D.S.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC097624; AAH97624.1; -; mRNA.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 595 AA; 65314 MW; 06883CF91309D29F CRC64;
Query Match 47.3%; Score 1431.5; DB 2; Length 595;
Best Local Similarity 47.7%; Pred. No. 2.1e-99;
Matches 277; Conservative 86; Mismatches 165; Indels 53; Gaps 6;
QY 21 AAKFHDVLGNRPSAYMRHNLGWSDDNDWNEKLYPVWKRGMKNSWKGRVQA 80
DB 19 AVKFPDVLGROPVGNRSHRHLGWSDDNDWNEKLYPVWKRGMKNSWKGRVQA 78
QY 81 VLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLSADPVYNNWNTAN 140
DB 79 RLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLSADPVYNNWNTAN 136
QY 141 LDDGNGTQSHNVFPDGPFPHPGWRNFIYVFTLGOYFQKLGRCSVRVSVNTAN 200
DB 137 -NDFCNGNCSFNKSPDGPFPHPGWRNFIYVFTLGOYFQKLGRCSVRVSVNTAN 195
QY 201 VTLPQLMETVYRRHGRVAVPIAQVDVYVVTQIPVFTVMFOKNDNSDDETLKDL 260
DB 196 ITAQVQVSVYRRHGRVAVPIAQVDVYVVTQIPVFTVMFOKNDNSDDETLKDL 255
QY 261 IMPDVLHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNLGDDTSLALSTLISVP 320
DB 256 INFDIRHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNLGDDTSLALSTLISVP 315
QY 321 GPCPP-----PPPPRPSKTP-----SLGAGDNPLESRIP----- 353
DB 316 SPCNPLATTPVTTFVTPVTTAQPMPFTTTSPTAQNTTGNSTDEPALFTTEVPTEITTT 375
QY 354 -----DENCQINRYGHQFQATITVEGILEVNIQMTDVLMPVWPRESSLIDFV 394
DB 376 TARTTTTARTTTTAAHTTAAAGCFIYRYGYSTTNTVVDGIVGIVEINIEMTNVQVPASQAE 435


```

QY 61 VVKGRDMRWKSNKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC 120
Db 61 VVKGRDMRWKSNKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC 120
QY 121 RNEAGLSADPYVNNWTANSESDGNGTQSHNVFPDGKPPHPGRRWNFIYVHTL 180
Db 121 RNEAGLSADPYVNNWTANSESDGNGTQSHNVFPDGKPPHPGRRWNFIYVHTL 180
QY 181 GOYFQ 185
Db 181 DWLLQ 185

RESULT 10
Q4RKH9 TETNG
ID Q4RKH9 TETNG PRELIMINARY; PRT; 538 AA.
AC Q4RKH9
DT 13-SEP-2005 (TremBLrel. 31, Created)
DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)
DE Chromosome 21 SCAP15029, whole genome shotgun sequence.
GN ORFNames=GSTENG00032949001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Crauad C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01015029; CAG11103.1; -; Genomic DNA.
SQ SEQUENCE 538 AA; 60051 MW; EAAALB2548813D9E CRC64;

Query Match 27.9%; Score 843; DB 2; Length 538;
Best Local Similarity 36.0%; Pred. No. 5.5e-55;
Matches 191; Conservative 84; Mismatches 214; Indels 42; Gaps 11;

QY 46 GWSNDENWEKLYPV--WKQDM-RWNSWKGVRVQAVLTSDSPALVGSNITFAVNLIF 102
Db 46 GWSNDENWEKLYPV--WKQDM-RWNSWKGVRVQAVLTSDSPALVGSNITFAVNLIF 102
QY 20 GWDQDTPWDDYLYPPLNLKTNELTRHK-----GKPKVHLTSDSPALVGSNITFAVNLIF 74
Db 20 GWDQDTPWDDYLYPPLNLKTNELTRHK-----GKPKVHLTSDSPALVGSNITFAVNLIF 74
QY 103 PRCKEDANGNIYKNCNEAGLSADPYVNNWTANSESDGNGTQSHNVFPDGKPP 162
Db 103 PRCKEDANGNIYKNCNEAGLSADPYVNNWTANSESDGNGTQSHNVFPDGKPP 162
QY 75 PPCQEDAGGDDVWDEHEDANGQIRSGVYNNWTANSESDGNGTQSHNVFPDGKPP 134
Db 75 PPCQEDAGGDDVWDEHEDANGQIRSGVYNNWTANSESDGNGTQSHNVFPDGKPP 134
QY 163 PHHPGRRWNFIYVHTLGOYFQKLGRCGRSVSVANTVTLGQPMLEVTYR-RHGRAYV 221
Db 163 PHHPGRRWNFIYVHTLGOYFQKLGRCGRSVSVANTVTLGQPMLEVTYR-RHGRAYV 221
QY 135 PQSNDWRHKSYYVWHTWQYETCDGSSSVYVNTHTPIGAEIMEVLVYKRRKYS 194
Db 135 PQSNDWRHKSYYVWHTWQYETCDGSSSVYVNTHTPIGAEIMEVLVYKRRKYS 194
QY 222 PTAQVKDYVVTDDQIPVFTVYVQKNDNRSSDFTFLKDLPIPFVLIHDPHSFL-NYSTIN 280
Db 222 PTAQVKDYVVTDDQIPVFTVYVQKNDNRSSDFTFLKDLPIPFVLIHDPHSFL-NYSTIN 280

```

```

Db 195 PLTVNTVYVYTDKIPAVASISQKAAVNSQSENVFPRGKDVVKQLHDPDGVLTAAALD 254
QY 281 YKWSFGDNTGLFVSTNHTVNTYVNLNGTSLNLTAKAAAGCPGPPPPP-----PPRPSK 334
Db 255 YIWDPRDGNQLVTHRNVT--THYSRLGTMSVKLVVEAAFPABCPPTSAPATERLPPTVTS 313
QY 335 PTPSLGPAGDNPLESRIIPDENQCINRCHQOATITIVEGILEVNIQWTDVLMVPWPPE 394
Db 314 ETPESGPTTLDLWLRVRLRSRQCFRYVHGTFVGNITIEPPFLKSA-----PH 362
QY 395 SSLID-----FVWTCOGSIPTEVCTIISDPTCEITONTVCSVDVDEMCLLTV 442
Db 363 SRIVDVASRVTKTDVSLVLCGLNIPVSACTIVSDPCTEVRAITCDDVPPLSCEVRL 422
QY 443 RRTFNGSGTYCVNLTLGDDTSLATSTLISV-PDRD-PASPLRMANSALISVGLAIFVT 500
Db 423 QRSFAEPCTYCVNITLEDGSGVALASTVITINKSDAPETSSRAAALSSAAVLWVFA 482
QY 501 VISLLVYKHKHYNPIENSPGNVRSKLSVFLNRAKAVFPFGNOEKDPLL 551
Db 483 FAALVLCRRYKGYRVPORPALGSCICRYGVGGVGHVHLREKLPFPRESHILL 533

RESULT 11
Q4KLIV9 XENLA
ID Q4KLIV9 XENLA PRELIMINARY; PRT; 719 AA.
AC Q4KLIV9
DT 13-SEP-2005 (TremBLrel. 31, Created)
DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)
DE LOC445868 protein (Fragment).
GN Name=LOC445868;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.C., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;

```


Klein S., Gerhard D.S.;
 Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC098976.1; -, mRNA.
 FT NON TER 1
 SQ SEQUENCE 719 AA; 76830 MW; 54A4D018A028C9E3 CRC64;

Query Match 24.4%; Score 736.5; DB 2; Length 719;
 Best Local Similarity 26.4%; Pred. No. 6.4e-47;
 Matches 189; Conservative 97; Mismatches 206; Indels 225; Gaps 19;

QY 43 QLINGWSSDE-----NDWNEKLYPVWKGDMRWKNSKGGRRVQAVLTSDS 86
 Db 21 QTONWSQNRVQTNQVQAGNROSFPKFSMSRMYPIWRGSEAQKNCWKGQVTFDLVND 80
 QY 87 PALVGSNITFAVNLIPRCQKEDANGNIYKENCNEAG--LSADPYVYNNTAWSESDG 144
 Db 81 PTLTGAKAFSILNFPNNQTVLPDQVWQGN-RTDNGTWIPSEPIYP-----DESTEG 135
 QY 145 ENTGOSHNVPPDGKPPPHHPGRRRNWPIYFHTLGOYFQKLGRCRSVRVSVNTANVTLG 204
 Db 136 SECT-----FPDGRPFPRGVEKHSKFVYVQWTKYQVVDGPPSNLTVETDGIPLG 188
 QY 205 PQLMMEVTVRRGR--AYVPIAQVKVYVTDQIPVFTVMFQKNDRNSDEFELKDLPIMF 263
 Db 189 SYTMQVVVHYRGQRKFIPIGISISQFTTIDQIPVSVSISQLLDQDQDFQINRAVSF 248
 QY 264 DVLHDPSPFLNYSNTYKWSFGDNTGLFVSNHTVNTHTVNLVNGFSLNLTAKAAP--GP 322
 Db 249 AVAIHDPHYLOADISFSWDFGDSGLTITRTDVTHTYVSPGVRPKVVLQAAIPIAP 308
 QY 323 C-----PPPPPPPRPSKTPPSLG-----PAG----- 343
 Db 309 CGSTAPVATAEPVPTTVAPAOPTTAAASTSGGVTDAPPGTTAALPGNITEPQGTINGIV 368
 QY 344 -----DN-----PLELSRPDE----- 355
 Db 369 VTIPSDSQNLAAEATLSSDNEAATLPEAIEDEAGTVVAGEETVPEEBAVFNQEQAVTV 428
 QY 356 ----- 355
 Db 429 AEAVPQEQAVTVNQEQAQVVAEAVPSQEAETSEATTETAIKLAEGSPSPVPSQINE 488
 QY 356 -----NCQINRYGHFOATITIVEGILEVNI 381
 Db 489 LPGTAAEVLTVPATAEAEVVIKRAQEPDPLVGLLYRYGTFTATDLDIQVIESAQIV 548
 QY 382 QMTDVLMPVPWPESSLLDFVTCQGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLT 441
 Db 549 QVAPV--AAVDGAENA-VDLTITCQGSVPQVCTIISNPDCTIPQSTVNCVTPQSTDCOLV 606
 QY 442 VRTENGSGYCVNLTLGDGDTSLATSLTSLISVPDRDPASPLRMANSALISVG-CLAIPTV 500
 Db 607 LRQVFNETGLYCVNVSILTDSVLSMASTQVSVSGAGSS-----FSGGGIVIVGVLLAVPAV 662
 QY 501 VISLLVYKHKHEYNPIEN--SPGNVVRSK-GLSVFLNRAKAVFFPGNCEKDPKLLKNQ 554
 Db 663 VIYAVTYRQKSKYALTAPSPSNWFPDRLSLRFLFQNALGLSRSG--ENSPLNGR 717

RESULT 12
 Q6DDN6 XENLA ID Q6DDN6 XENLA PRELIMINARY; PRT; 721 AA.
 AC Q6DDN6
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE LOC445868 protein (Fragment).
 GN Names=LOC445868;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC xenopodinae; Xenopus; Xenopus.
 NCBI TaxId=8355;

[illegible]

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RP [3]
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature04266;
 RC Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaado I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matusda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Mulcahy L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynehaw-Boris A., Yangisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RL 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RP [4]
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RC Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RP [5]
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RC Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hachima M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Hashiwa K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RL sequencing pipeline with 384 multiplex capillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RP [6]
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito K., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,

```
Qy 323 CPPPPP-----PPPS-----K 334
Db | | | | |
339 MPTTQPSGTTVVQMPTEVTATTSEQMLTSAVIDTTLAENVSTTEGTGTTPTRPSTGTTVAQ 398
Qy 335 PTPSLGPAGDNPL-----ELSRIPDE-----NCQINRYGHFQATITIVE 373
Db | | | | |
399 ATTTEGPDASPLLPSTOSSTGSI SPLLD DDTIMLVKRQVPLDCVLYRYGSFSLALDIVQ 457
Qy 374 GILEVNIIQMTDVLMPVWPMPPESSLIDFVVTCCGSIPTTEVCTIISDPTCEITQNTVCSPVD 433
Db | | | | |
458 G-----IESAEIILQAVPFSEGDAPFELTVSCQGLPKXACMDISSPGCQPPAORLCOSVP 511
Qy 434 VDEMCLLTVRRTF-NGSGTYCVNLTGLGDDTSLALTSTLISVPDRDPASPLRMANSALISV 492
Db | | | | |
512 PSPDCQLVHLQVLKGGSGTYCLNVSLADANSLAVASTQLVVPQDGG-----LCQAPLLV 566
Qy 493 GCLAIFTVTVISLLVYKXH--KEYNPIENSPGN-----VVRSGLSVFLNRAKAVFF 541
Db | | | | |
567 GILLVLAVAVLASLIHRRLKKGQSVSQMPHGSTHMLRLPVFRARGIG----- 615
Qy 542 PGNOEKDPLKNOE 555
Db | | | | |
616 -----ENSPLLSGQQ 625
```

Search completed: December 24, 2005, 04:35:47
Job time : 243 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:28:24 ; Search time 49 Seconds
(without alignments)
944.865 Million cell updates/sec

Title: US-09-853-880A-17
Perfect score: 3026
Sequence: 1 MECLYFGLPILLARLPD.....PFGNQKDPLLNQKFGKVS 560

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	1 US-08-594-031-90	Sequence 90, Appl
2	3026	100.0	560	2 US-09-643-597-225	Sequence 225, App
3	3026	100.0	560	2 US-09-480-884A-225	Sequence 225, App
4	3026	100.0	560	2 US-09-542-615A-225	Sequence 225, App
5	3026	100.0	560	2 US-09-606-421B-225	Sequence 225, App
6	3026	100.0	560	2 US-09-476-496A-225	Sequence 225, App
7	3026	100.0	560	2 US-09-630-940B-225	Sequence 225, App
8	3026	100.0	560	2 US-09-943-075A-6	Sequence 6, Appl
9	3026	100.0	560	2 US-10-039-272A-2	Sequence 2, Appl
10	3026	100.0	560	2 US-10-007-700-225	Sequence 225, App
11	3026	100.0	560	5 US-09-985-799-90	Sequence 90, Appl
12	3026	100.0	560	5 US-09-977-371-90	Sequence 36, Appl
13	2197	72.6	574	2 US-09-383-586-36	Sequence 36, Appl
14	2197	72.6	574	2 US-09-823-038A-36	Sequence 36, Appl
15	2197	72.6	574	2 US-09-943-075A-5	Sequence 5, Appl
16	2179	72.0	572	2 US-09-197-970B-5	Sequence 2, Appl
17	2163	71.5	572	2 US-09-943-075A-2	Sequence 2, Appl
18	969	32.0	376	1 US-08-594-031-100	Sequence 100, App
19	969	32.0	376	1 US-08-594-031-102	Sequence 102, App
20	969	32.0	376	5 US-09-985-799-100	Sequence 100, App
21	969	32.0	376	5 US-09-985-799-102	Sequence 102, App
22	969	32.0	376	5 US-09-977-371-100	Sequence 100, App
23	969	32.0	376	5 US-09-977-371-102	Sequence 102, App
24	622	20.6	661	1 US-08-417-174-121	Sequence 121, App
25	622	20.6	661	2 US-09-267-439-121	Sequence 121, App
26	622	20.6	661	2 US-08-388-852B-2	Sequence 2, Appl
27	622	20.6	661	2 US-09-073-138-121	Sequence 121, App

Query Match 100.0%; Score 3026; DB 1; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;

28	622	20.6	661	2 US-09-862-260A-2	Sequence 2, Appl
29	621	20.5	661	1 US-08-417-174-27	Sequence 27, Appl
30	621	20.5	661	1 US-08-231-565A-27	Sequence 27, Appl
31	621	20.5	661	1 US-09-007-961-27	Sequence 27, Appl
32	621	20.5	661	2 US-09-267-439-27	Sequence 27, Appl
33	621	20.5	661	2 US-09-073-138-27	Sequence 27, Appl
34	603.5	19.9	668	1 US-07-891-942G-6	Sequence 6, Appl
35	354.5	11.7	460	2 US-09-949-016-8029	Sequence 8029, Ap
36	238.5	7.9	190	2 US-08-388-852B-35	Sequence 35, Appl
37	236	7.8	192	2 US-08-388-852B-38	Sequence 38, Appl
38	233.5	7.7	192	2 US-08-388-852B-36	Sequence 36, Appl
39	233	7.7	197	2 US-08-388-852B-37	Sequence 37, Appl
40	181.5	6.0	423	2 US-09-991-181-241	Sequence 241, App
41	181.5	6.0	423	2 US-09-990-444-241	Sequence 241, App
42	181.5	6.0	423	2 US-09-997-333-241	Sequence 241, App
43	181.5	6.0	423	2 US-09-992-598-241	Sequence 241, App
44	136.5	4.5	71	1 US-08-594-031-98	Sequence 98, Appl
45	136.5	4.5	71	5 US-09-985-799-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1
US-08-594-031-90
; Sequence 90, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE: 30-JAN-1996
; APPLICATION NUMBER: US/08/594,031
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-594-031-90

Query Match 100.0%; Score 3026; DB 1; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;

1	MECLYYTGLFLLAARLPDIDAAKGFHDVIGNERPSAYNRHQNQNGWSSDENDWNEKLP	60
61	VVKRGDMRWKNSWKGRYQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYKNC	120
61	VVKRGDMRWKNSWKGRYQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYKNC	120
121	RNEAGLSADPYYNNTAWSESDSDENGTGQSHHNVFPDGKPPPHHPCGRRRNPIYVFHTL	180
121	RNEAGLSADPYYNNTAWSESDSDENGTGQSHHNVFPDGKPPPHHPCGRRRNPIYVFHTL	180
181	GOYFOKJGRCSVRSVNTANVTGLPGQLMVTVYRRHGRAVVPYTAQVKDQVYVTVTDQIPV	240
181	GOYFOKJGRCSVRSVNTANVTGLPGQLMVTVYRRHGRAVVPYTAQVKDQVYVTVTDQIPV	240
241	TMFQKNDNRSSDETEFLKDLPIPMFVDLIHDPHSFLNYSTINYKWSFGDNTGLFVSTNHTVN	300
241	TMFQKNDNRSSDETEFLKDLPIPMFVDLIHDPHSFLNYSTINYKWSFGDNTGLFVSTNHTVN	300
301	HTYVLNGTFSMLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN	360
301	HTYVLNGTFSMLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN	360
361	RYGHQATITIVEGILEVNIIQMTDVLMPVPWPRESSLIIDFWVTCQSGIPEVCTIISDPT	420
361	RYGHQATITIVEGILEVNIIQMTDVLMPVPWPRESSLIIDFWVTCQSGIPEVCTIISDPT	420
421	CSBITQNTVCSVPDVBDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLATSTLIISVPDRDRPAS	480
421	CSBITQNTVCSVPDVBDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLATSTLIISVPDRDRPAS	480
481	PLRMANSALISVGCILAI FVTYISLLVYKXKHEYNPIENSPGNVVRSGKLSVFLNRAKAYP	540
481	PLRMANSALISVGCILAI FVTYISLLVYKXKHEYNPIENSPGNVVRSGKLSVFLNRAKAYP	540
541	FPGNQEKDPLLNQBFKGV	560
541	FPGNQEKDPLLNQBFKGV	560

RESULT 4

US-09-542-615A-225
; Sequence 225, Application US/09542615A

Patent No. 6518256

```

1  / GENERAL INFORMATION:
2  /
3  / APPLICANT: Wang, Tongtong
4  / APPLICANT: Fan, Liqun
5  / APPLICANT: Kalos, Michael D.
6  / APPLICANT: Bangur, Chaitanya S.
7  / APPLICANT: Hosken, Nancy A.
8  / APPLICANT: Fanger, Gary R.
9  /
10 / TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
11 /
12 / FILE REFERENCE: 210121.455C8
13 /
14 / CURRENT APPLICATION NUMBER: US/09/542,615A
15 /
16 / CURRENT FILING DATE: 2000-04-14
17 /
18 / NUMBER OF SEQ ID NOS: 350
19 /
20 / SOFTWARE: FastSeq for Windows Version 3.0
21 /
22 / SEQ ID NO 225
23 /
24 / LENGTH: 560
25 /
26 / TYPE: prt
27 /
28 / ORGANISM: Homo sapien
29 /

```

```
Query Match      100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1	MECLYYFLGFLLLAARLPDLDAAKRPHDVLGNERSAYMREHNLQNGWSSDENDWNEKLYP	60
Qy		
1	MECLYYFLGFLLLAARLPDLDAAKRPHDVLGNERSAYMREHNLQNGWSSDENDWNEKLYP	60
Db		
61	VMKRGDMRWKNSWKGRRVOAVLTDSPALVGSNTTFAVNLIFFRCQKEDANGNIVYEKNC	120
Qy		

61	WVRGDMRWKQXWKGGRVQAVLTSDSPALVGSNTTPAVNLIPPRCQKEDANGNIYVKVC	120
121	RNEAGLSADPYVYNWTAWSEDSDENGTQSHHNVPDGGKPFPHHPGWRWNFIYVPHLL	180
121	RNEAGLSADPYVYNWTAWSEDSDENGTQSHHNVPDGGKPFPHHPGWRWNFIYVPHLL	180
181	GOYFOKLGRCSVRSVNNTANVTLGQPMLEVVTVYRRHGRAYVPIAQVKDVVVVTDQIPVVF	240
181	GOYFOKLGRCSVRSVNNTANVTLGQPMLEVVTVYRRHGRAYVPIAQVKDVVVVTDQIPVVF	240
241	TMEQKQDRNSSDETFLKDLPIMEDVLTHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN	300
241	TMEQKQDRNSSDETFLKDLPIMEDVLTHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN	300
301	HTVVLNGTFSNLNTVKAAAAGPCPPPPPPRRPSKPTPSLGPAGNPJLSRIPDENCQIN	360
301	HTVVLNGTFSNLNTVKAAAAGPCPPPPPPRRPSKPTPSLGPAGNPJLSRIPDENCQIN	360
361	RYGHFQATITIVBGIILEVNIIQMTDVLMPVPWPRESSLIIDFVVTCQGSIPTEVCITISDPT	420
361	RYGHFQATITIVBGIILEVNIIQMTDVLMPVPWPRESSLIIDFVVTCQGSIPTEVCITISDPT	420
421	CEITQNTVCSPVDVDECLLITVTRATFNGSGTYCVNLTLGGDTSIALSTLSLSVDPDRDPAS	480
421	CEITQNTVCSPVDVDECLLITVTRATFNGSGTYCVNLTLGGDTSIALSTLSLSVDPDRDPAS	480
481	PLRMANSALITSVGCLAI FVTVVISLVYVKKHCKEYNPIENSQGNVVRSGLSVFLNRAKAVF	540
481	PLRMANSALITSVGCLAI FVTVVISLVYVKKHCKEYNPIENSQGNVVRSGLSVFLNRAKAVF	540
541	FPNGQEKDPLLKNQEFKGVSS	560
541	FPNGQEKDPLLKNQEFKGVSS	560

RESULT 5

US-09-606-421B-225
; Sequence 225, Application US/09606421B

; Patent No. 6531315

```

1  GENERAL INFORMATION:
2  APPLICANT: Wang, Tongtong
3  APPLICANT: Fan, Lique
4  APPLICANT: Kalos, Michael D.
5  APPLICANT: Bangur, Chaitanya S.
6  APPLICANT: Hosken, Nancy
7  APPLICANT: Panger, Gary R.
8  APPLICANT: Li, Samuel X.
9  APPLICANT: Wang, Aijun
10 APPLICANT: Skeiky, Yasir A.W.
11 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
12 OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
13 FILE REFERENCE: 210121.455C9
14 CURRENT APPLICATION NUMBER: US/09/606,421B
15 CURRENT FILING DATE: 2000-06-28
16 NUMBER OF SEQ ID NOS: 358
17 SOFTWARE: FastSEQ for Windows Version 3.0
18 SEQ ID NO 225
19 LENGTH: 560
20 TYPE: PRT
21 ORGANISM: Homo sapien
22 US-09-606-421B-225

```

```
Query Match      100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

1	MECLYYFLGFLLLAARLPDLAAKGFHDVIGNERPSAYMREHNQINGWSSDSDNDWNEKLYP	60
Qy		
1	MECLYYFLGFLLLAARLPDLAAKGFHDVIGNERPSAYMREHNQINGWSSDSDNDWNEKLYP	60
Db		
61	VWKRGMRWKSNWKGGRVQAVLTSDSPALVGSNTTFANLIFPRCQKEDANGNTVYEKNC	120
Qy		
61	VWKRGMRWKSNWKGGRVQAVLTSDSPALVGSNTTFANLIFPRCQKEDANGNTVYEKNC	120
Db		

QY 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180
QY 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240
DB 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240
QY 241 TMFQKNDNSDDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSDDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNLTAKAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
DB 301 HTYVLNGTFSNLNLTAKAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
QY 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPBESSLIDFVTCQGSIPTEVCTIISDPT 420
DB 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPBESSLIDFVTCQGSIPTEVCTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLSVDPDRPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLSVDPDRPAS 480
QY 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540
QY 541 FPGNQEKDPLLNKQEFKGS 560
DB 541 FPGNQEKDPLLNKQEFKGS 560

RESULT 6

US-09-476-496A-225
; Sequence 225, Application US/09476496A
; Patent No. 6706262

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: LUNG CANCER
; FILE REFERENCE: 210121.455C5
; CURRENT APPLICATION NUMBER: US/09/476.496A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-476-496A-225

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180
QY 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240

DB 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240
QY 241 TMFQKNDNSDDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSDDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNLTAKAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
DB 301 HTYVLNGTFSNLNLTAKAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
QY 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPBESSLIDFVTCQGSIPTEVCTIISDPT 420
DB 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPBESSLIDFVTCQGSIPTEVCTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLSVDPDRPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLSVDPDRPAS 480
QY 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540
QY 541 FPGNQEKDPLLNKQEFKGS 560
DB 541 FPGNQEKDPLLNKQEFKGS 560

RESULT 7

US-09-630-940B-225
; Sequence 225, Application US/09630940B
; Patent No. 6737514

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630.940B
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-630-940B-225

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGKPPPHGWRWNFIYVFTL 180

QY 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
QY 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360
QY 361 RYGHFOARTITIVEGILEVNIQMTDVLMPVPMPPESSLIDFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOARTITIVEGILEVNIQMTDVLMPVPMPPESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVPDVEDMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
DB 421 CEITQNTVCSVPDVEDMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGLCLAIPTVVISLVYKHKKEYNPPIENSPGNVVRSGKLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGLCLAIPTVVISLVYKHKKEYNPPIENSPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQBFKGV 560
DB 541 PFGNQEKDPLLKNQBFKGV 560

RESULT 8

US-09-943-075A-6
; Sequence 6, Application US/09943075A
; Patent No. 6812002
; GENERAL INFORMATION:
; APPLICANT: Popoff, Steven N.
; APPLICANT: Safado, Faye P.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,
; FILE REFERENCE: 71369.262
; CURRENT APPLICATION NUMBER: US/09/943,075A
; PRIOR APPLICATION NUMBER: US 60/229,006
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Human
US-09-943-075A-6

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFFGLFLLAARLPDAAKRFHDLGNRPSAYMRHNLQNGSSDENDWNEKLYP 60
DB 1 MECLYFFGLFLLAARLPDAAKRFHDLGNRPSAYMRHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVNTWTAMSESDGNGTQSHHNPFGKPPHHPGWRWRNFIYVFTL 180
DB 121 RNEAGLSADPYVNTWTAMSESDGNGTQSHHNPFGKPPHHPGWRWRNFIYVFTL 180
QY 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240

QY 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360
QY 361 RYGHFOARTITIVEGILEVNIQMTDVLMPVPMPPESSLIDFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOARTITIVEGILEVNIQMTDVLMPVPMPPESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVPDVEDMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
DB 421 CEITQNTVCSVPDVEDMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGLCLAIPTVVISLVYKHKKEYNPPIENSPGNVVRSGKLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGLCLAIPTVVISLVYKHKKEYNPPIENSPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQBFKGV 560
DB 541 PFGNQEKDPLLKNQBFKGV 560

RESULT 9

US-10-039-272A-2
; Sequence 2, Application US/10039272A
; Patent No. 6939955
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE
; FILE REFERENCE: 267/033 (UMD-0055)
; CURRENT APPLICATION NUMBER: US/10/039,272A
; CURRENT FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-039-272A-2

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFFGLFLLAARLPDAAKRFHDLGNRPSAYMRHNLQNGSSDENDWNEKLYP 60
DB 1 MECLYFFGLFLLAARLPDAAKRFHDLGNRPSAYMRHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVNTWTAMSESDGNGTQSHHNPFGKPPHHPGWRWRNFIYVFTL 180
DB 121 RNEAGLSADPYVNTWTAMSESDGNGTQSHHNPFGKPPHHPGWRWRNFIYVFTL 180
QY 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
QY 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSDETFLKDLPIMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGAGDNPLSLSRIPDENCCQIN 360

QY 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPRESSLIDFVVTQCGSIPTVEVCTIIISDPT 420
DB 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPRESSLIDFVVTQCGSIPTVEVCTIIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
QY 481 PLRMANSLISVGLAIFVTVISLLVYKHKKEYNPNIENSPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSLISVGLAIFVTVISLLVYKHKKEYNPNIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PPGNQEKDPLLNQEFKGV 560
DB 541 PPGNQEKDPLLNQEFKGV 560

RESULT 10
US-10-007-700-225
; Sequence 225, Application US/10007700
; Patent No. 6960570
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-225

Query Match 100.0%; Score 3026; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKPHDVLGNRPISAYNRHNLNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKPHDVLGNRPISAYNRHNLNGWSSDENDWNEKLYP 60
QY 61 VWRGDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIVYEKNC 120
DB 61 VWRGDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIVYEKNC 120
QY 121 RNEAGLSADPVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHHPGWRRWNFIYVFTTL 180
DB 121 RNEAGLSADPVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHHPGWRRWNFIYVFTTL 180
QY 181 GQYFQKLGRCVRSVNTANTVLTGQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANTVLTGQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240
QY 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300

DB 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTVYVINGTFTSLNTVKAAPGCPGPPPPRPSKPTPSLPGAGDNPLELSRIPDENCOIN 360
DB 301 HTVYVINGTFTSLNTVKAAPGCPGPPPPRPSKPTPSLPGAGDNPLELSRIPDENCOIN 360
QY 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPRESSLIDFVVTQCGSIPTVEVCTIIISDPT 420
DB 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPRESSLIDFVVTQCGSIPTVEVCTIIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480
QY 481 PLRMANSLISVGLAIFVTVISLLVYKHKKEYNPNIENSPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSLISVGLAIFVTVISLLVYKHKKEYNPNIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PPGNQEKDPLLNQEFKGV 560
DB 541 PPGNQEKDPLLNQEFKGV 560

RESULT 11
US-09-985-799-90
; Sequence 90, Application US/09985799
; Patent No. RE38392
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/985,799
; FILING DATE: 06-No. RE38392-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:

US-09-985-799-90

Query Match 100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVNTAWSESDGNGTQSHHNVFPDGPDPHHPGWRWNFIYVFHTL 180
DB 121 RNEAGLSADPYVNTAWSESDGNGTQSHHNVFPDGPDPHHPGWRWNFIYVFHTL 180
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVVTDOIIPVF 240
DB 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVVTDOIIPVF 240
QY 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDNCQIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDNCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPRESSLIDFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPRESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITONTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480
DB 421 CEITONTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGCIAIFVTVISLAVYKHKKEYNPINSPGNVVRSGKLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCIAIFVTVISLAVYKHKKEYNPINSPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQEFKGV 560
DB 541 PFGNQEKDPLLNQEFKGV 560

RESULT 12

US-09-977-371-90

; Sequence 90, Application US/09977371

; Patent No. RE38490

; GENERAL INFORMATION:

; APPLICANT: THOMPSON, Timothy C.

; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES

; NUMBER OF SEQUENCES: 175

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BAKER & BOTTS, L.L.P.

; STREET: 1299 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20004-2400

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/977,371

; FILING DATE: 16-Oct-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/594,031

; APPLICATION NUMBER: US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; FILING DATE: 30-JAN-1996
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:
; US-09-977-371-90

Query Match 100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.6e-280;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVNTAWSESDGNGTQSHHNVFPDGPDPHHPGWRWNFIYVFHTL 180
DB 121 RNEAGLSADPYVNTAWSESDGNGTQSHHNVFPDGPDPHHPGWRWNFIYVFHTL 180
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVVTDOIIPVF 240
DB 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVVTDOIIPVF 240
QY 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDNCQIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDNCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPRESSLIDFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPMPRESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITONTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480
DB 421 CEITONTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGCIAIFVTVISLAVYKHKKEYNPINSPGNVVRSGKLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCIAIFVTVISLAVYKHKKEYNPINSPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQEFKGV 560
DB 541 PFGNQEKDPLLNQEFKGV 560

RESULT 13

US-09-983-586-36

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586

; Sequence 36, Application US/0983586


```
FILE REFERENCE: 71369.262
CURRENT APPLICATION NUMBER: US/09/943,075A
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: US 60/229,006
PRIOR FILING DATE: 2000-08-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 574
TYPE: PRT
ORGANISM: Mouse
US-09-943-075A-5

Query Match      72.6%; Score 2197; DB 2; Length 574;
Best Local Similarity 70.7%; Pred. No. 3.3e-201;
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

QY      1 MESLYYFGLFLLAARLPDAAKRFHDVLGNRPSPAYMREHNLQNGSSDENDWNEKLYP 60
DB      1 MESLGVGLFLLAARLPDAAKRFHDVLGNRPSPAYMREHNLQNGSSDENDWNEHLYP 60

QY      61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITPAVNLIFFRCQKEDANGNIYVEKNC 120
DB      61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITPAVNLIFFRCQKEDANGNIYVEKNC 120

QY      121 RNEAGLSADPYNTWTAWSESDGNGTCQSHNVFPDGKPPHPGWRNPNFYVFTYL 180
DB      121 RNDLGLTSLHVNNTAGADGDWEDGTSRQHLRFPDRRPPRPHGKWKNSFVYVFTYL 180

QY      181 GQYFQKLGRCVSVRVNTANTVLGQPMVTVYRRHGRAYVPIAOVKVYVVTDOIPIV 240
DB      181 GQYFQKLGRCVSVRVNTANTVLGQPMVTVYRRHGRAYVPIAOVKVYVVTDOIPIV 240

QY      241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPHFHLYSTINYKWSFGDNTGLFVSTNHTVN 300
DB      241 TMSQKNDNRNLSDIIFLDLPVDFVLIHDPHFHLYSTINYKWSFGDNTGLFVSNHTLN 300

QY      301 HTYVLNGTFPSNLTVKAAAPGPCPPPP--PPRP-----SKPTPSLGPAGD 344
DB      301 HTYVLNGTFNLTVQTAVPGPCPPSPSTPPSPSTPPSPPLTSTPSPLMPTGY 360

QY      345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVWPBESSLIDFVTC 404
DB      361 KSMELSDISNENCRINRYGFRTATITVEGILEVSIQADVPMPPTQPANSLMDFTVTC 420

QY      405 QGSIPTEVCTIISDPTCEITQNTVCSPVDVDEMCLITVARTFNGSGTYCVNLTGDDTSL 464
DB      421 KGATPWEACTIISDPTCQIAQNRVCSVAVDGLCLLSVERAFNGSGTYCVNLTGDDASL 480

QY      465 ALTSTLISVPDRDPASPLRMANSALISVGCLAI FVTVISLIVYKKHKEYNP IENSPGNV 524
DB      481 ALTSTLISIPGKDPSPRAVNGVLISIGCLAVLVTMTVITLLYKKHKEYKPIGNCPRNTV 540

QY      525 RSKGLSVFLNRAKAVFPFCNQEKDPLLNQ 554
DB      541 KKGGLSVLLSHAKAPFRGDKQKDLQDK 570
```

Search completed: December 24, 2005, 04:37:24
Job time : 52 secs

THIS PAGE BLANK (USPTO)


```
Qy 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
Db 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
Qy 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRTTFNGSGTYCNVLTGLGDDTSLALTSTLISVDDRDPAS 480
Db 421 CEITQNTVCSVDVDEMCLLTVRTTFNGSGTYCNVLTGLGDDTSLALTSTLISVDDRDPAS 480
Qy 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560

RESULT 2
US-09-850-716A-225
; Sequence 225, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retex, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-225
```

```
Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGLLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Db 1 MECLYYFLGLLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Qy 61 VKRGMWRKNSWKGRRVQAVLTSDSPALVGSNTITFAVNIIFPRCKEDANGNIYVYKNC 120
Db 61 VKRGMWRKNSWKGRRVQAVLTSDSPALVGSNTITFAVNIIFPRCKEDANGNIYVYKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHHNVFPDGKFPFHPHGWRRNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHHNVFPDGKFPFHPHGWRRNFIYVFHTL 180
Qy 181 GOYFQKLGRCRVS RVSNVTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPV 240
Db 181 GOYFQKLGRCRVS RVSNVTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPV 240
Qy 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
Db 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
```

```
Qy 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRTTFNGSGTYCNVLTGLGDDTSLALTSTLISVDDRDPAS 480
Db 421 CEITQNTVCSVDVDEMCLLTVRTTFNGSGTYCNVLTGLGDDTSLALTSTLISVDDRDPAS 480
Qy 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560

RESULT 3
US-09-897-778-225
; Sequence 225, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-225
```

```
Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGLLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Db 1 MECLYYFLGLLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Qy 61 VKRGMWRKNSWKGRRVQAVLTSDSPALVGSNTITFAVNIIFPRCKEDANGNIYVYKNC 120
Db 61 VKRGMWRKNSWKGRRVQAVLTSDSPALVGSNTITFAVNIIFPRCKEDANGNIYVYKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHHNVFPDGKFPFHPHGWRRNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHHNVFPDGKFPFHPHGWRRNFIYVFHTL 180
Qy 181 GOYFQKLGRCRVS RVSNVTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPV 240
Db 181 GOYFQKLGRCRVS RVSNVTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPV 240
Qy 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSEFLKDLPIIMFDVLIDHPSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
Db 301 HTYYVLNGTFSNLITVKAAPGCPGPPPPRPSKPTPSLGPAGDNPISLRIIPDENCQIN 360
```


QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
QY 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQEFKGS 560
Db 541 PFGNQEKDPLLNQEFKGS 560

RESULT 4

US-09-943-075A-6
; Sequence 6, Application US/09943075A
; Patent No. US20020151486A1
; GENERAL INFORMATION:
; APPLICANT: Popoff, Steven N.
; APPLICANT: Safado, Pavez P.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Osteocalcin Protein and Nucleic Acids Encoding the Same,
; FILE REFERENCE: 71369.262
; CURRENT APPLICATION NUMBER: US/09/943,075A
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,006
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Human
US-09-943-075A-6

Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMREHNQNLGWSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMREHNQNLGWSDENDWNEKLYP 60
QY 61 VKKGDWRKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VKKGDWRKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVYNTWNTAWSESDGNGTQSHHNVFPDCKPPHHPGRRWNFIYVFHTL 180
Db 121 RNEAGLSADPVYVYNTWNTAWSESDGNGTQSHHNVFPDCKPPHHPGRRWNFIYVFHTL 180
QY 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
Db 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
QY 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLSLRIPDENCQIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLSLRIPDENCQIN 360
QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420

QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
QY 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQEFKGS 560
Db 541 PFGNQEKDPLLNQEFKGS 560

RESULT 5

US-10-039-272-2
; Sequence 2, Application US/10039272
; Publication No. US20020168653A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/039,272
; CURRENT FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-039-272-2

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMREHNQNLGWSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMREHNQNLGWSDENDWNEKLYP 60
QY 61 VKKGDWRKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VKKGDWRKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVYNTWNTAWSESDGNGTQSHHNVFPDCKPPHHPGRRWNFIYVFHTL 180
Db 121 RNEAGLSADPVYVYNTWNTAWSESDGNGTQSHHNVFPDCKPPHHPGRRWNFIYVFHTL 180
QY 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
Db 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYVPIAQVKDYYVVTQDQIPVFV 240
QY 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLSLRIPDENCQIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLSLRIPDENCQIN 360
QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPRDRPAS 480
QY 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCCLAIPTVVISLIVYKHKKEYNPISPGNVVRSGKLSVFLNRAKAVF 540

Qy 541 FPGNQEKDPLLNQEFKGV 560
Db 541 FPGNQEKDPLLNQEFKGV 560

RESULT 6

US-10-007-700-225
; Sequence 225, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007.700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-225

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MECLYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFTL 180
Qy 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVVVTQDIPFV 240
Db 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVVVTQDIPFV 240
Qy 241 TMFQKNDNRNSDDETFKDLPIFMDVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIFMDVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPSKPTPSLGPAGNPLELSRIPDENCOIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPSKPTPSLGPAGNPLELSRIPDENCOIN 360
Qy 361 RYGHFQATITIVGILEVNIQMTDVLMPWPWPESSLIIDFVVTCCGSIPTVEVCTIISDPT 420
Db 361 RYGHFQATITIVGILEVNIQMTDVLMPWPWPESSLIIDFVVTCCGSIPTVEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSIALTSLTISVDRDPAS 480

Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSIALTSLTISVDRDPAS 480
Qy 481 PLRMANSALISVGCLAFVTVVISLLVYKHKENYNPISPGNVRSKGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAFVTVVISLLVYKHKENYNPISPGNVRSKGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNQEFKGV 560
Db 541 FPGNQEKDPLLNQEFKGV 560

RESULT 7

US-10-117-982-225
; Sequence 225, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117.982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-225

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MECLYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGSSDENDWNEKLYP 60
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFTL 180
Qy 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVVVTQDIPFV 240
Db 181 GOYFQKLGRCRSVRVSVNTANVTLPQMEVTVYRRHGRAYVPIAQKDVVVTQDIPFV 240
Qy 241 TMFQKNDNRNSDDETFKDLPIFMDVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIFMDVLIHDPGSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPSKPTPSLGPAGNPLELSRIPDENCOIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPPSKPTPSLGPAGNPLELSRIPDENCOIN 360
Qy 361 RYGHFQATITIVGILEVNIQMTDVLMPWPWPESSLIIDFVVTCCGSIPTVEVCTIISDPT 420
Db 361 RYGHFQATITIVGILEVNIQMTDVLMPWPWPESSLIIDFVVTCCGSIPTVEVCTIISDPT 420

QY 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRPAS 480
QY 481 PLRMANALISVGCLAIFFVTISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
Db 481 PLRMANALISVGCLAIFFVTISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQBFKGV 560
Db 541 PFGNQEKDPLLKNQBFKGV 560
RESULT 8
US-10-463-106-2
; Sequence 2, Application US/10463106
; Publication No. US20030202938A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/463,106
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 10/039,272
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-463-106-2
Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDGPFPHPGRRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDGPFPHPGRRWNFIYVFTL 180
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQLMVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
Db 181 GOYFQKGRCSVRVSVNTANVTGLPQLMVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240
QY 241 TMFQKNDNRSSDETFLKPLIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRSSDETFLKPLIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTSLNLTAKAAGCPPP 360
Db 301 HTYVLNGTSLNLTAKAAGCPPP 360
QY 361 RYGHFOATITVEGILEVNIQMTDVLMPVPPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db 361 RYGHFOATITVEGILEVNIQMTDVLMPVPPESSLIDFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRPAS 480
QY 481 PLRMANALISVGCLAIFFVTISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540

Db 481 PLRMANALISVGCLAIFFVTISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQBFKGV 560
Db 541 PFGNQEKDPLLKNQBFKGV 560
RESULT 9
US-10-295-027-1258
; Sequence 1258, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1258
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1258
Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDGPFPHPGRRWNFIYVFTL 180
Db 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDGPFPHPGRRWNFIYVFTL 180
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQLMVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240

Db 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVDVYVVTQIEPVF 240
Qy 241 TMFQKNDNRSSDTEFLKDLPTIMFDVLHDPESHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRSSDTEFLKDLPTIMFDVLHDPESHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLNLTVKAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Db 301 HTYVLNGTFSNLNLTVKAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Qy 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQGSIPTEVCTIISDPT 420
Db 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQGSIPTEVCTIISDPT 420
Qy 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480
Qy 481 PLRMANSALISVGCLAIFFTIVTISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIFFTIVTISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNQKQEFKGV 560
Db 541 FPGNQEKDPLLNQKQEFKGV 560

RESULT 10
US-10-313-986-225
; Sequence 225, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Tereesa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313.986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-225

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERSAYMRBHNQNGSSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERSAYMRBHNQNGSSDENDWNEKLYP 60
Qy 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYBKN 120
Db 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYBKN 120
Qy 121 RNEAGLSADPVYVNTWTAWSESDSGENGTSQSHNVFPDGKPFPHHGRWRNFIYVFHTL 180
Db 121 RNEAGLSADPVYVNTWTAWSESDSGENGTSQSHNVFPDGKPFPHHGRWRNFIYVFHTL 180
Qy 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVDVYVVTQIEPVF 240
Db 181 GQYFQKLGRCVRSVNTANVTLPQMEVTVYRRHGRAYVPIAQVDVYVVTQIEPVF 240
Qy 241 TMFQKNDNRSSDTEFLKDLPTIMFDVLHDPESHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300

Db 241 TMFQKNDNRSSDTEFLKDLPTIMFDVLHDPESHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYVLNGTFSNLNLTVKAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Db 301 HTYVLNGTFSNLNLTVKAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Qy 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQGSIPTEVCTIISDPT 420
Db 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQGSIPTEVCTIISDPT 420
Qy 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480
Qy 481 PLRMANSALISVGCLAIFFTIVTISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIFFTIVTISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNQKQEFKGV 560
Db 541 FPGNQEKDPLLNQKQEFKGV 560

RESULT 11
US-10-309-290-152
; Sequence 152, Application US/10309290
; Publication No. US20040023241A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chillakuru, Rajeev A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Gould-Rothberg, Bonnie E.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Jeffers, Michael E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Miller, Charles E.
; APPLICANT: Murphey, Ryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Peyman, John A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Smithson, Glennda
; APPLICANT: Starling, Gary
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Voess, Edward Z.
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-502A
; CURRENT APPLICATION NUMBER: US/10/309,290
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/336,600
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,285
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/341,346
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,297
; PRIOR FILING DATE: 2001-12-27

```
; PRIORITY APPLICATION NUMBER: 60/344,903
; PRIORITY FILING DATE: 2001-12-31
; PRIORITY APPLICATION NUMBER: 60/373,288
; PRIORITY FILING DATE: 2002-04-17
; PRIORITY APPLICATION NUMBER: 60/380,981
; PRIORITY FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 152
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-309-290-152

Query Match      100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
DB 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
QY 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
DB 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
QY 181 GQYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIPVFV 240
QY 241 TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTKAAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
DB 301 HTYVLNGTFSNLTKAAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLDIFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLDIFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480
QY 481 PLRMANSALISVGCCLAFVTVISLLVYKKHKEYNPISPGNVVRSGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCCLAFVTVISLLVYKKHKEYNPISPGNVVRSGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQKQFEGVS 560
DB 541 PFGNQEKDPLLNQKQFEGVS 560

RESULT 12
US-10-408-765A-466
; Sequence 466, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
```

```
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 466
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-466

Query Match      100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
DB 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
QY 181 GQYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIPVFV 240
DB 181 GQYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIPVFV 240
QY 241 TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTKAAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
DB 301 HTYVLNGTFSNLTKAAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLDIFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLDIFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480
QY 481 PLRMANSALISVGCCLAFVTVISLLVYKKHKEYNPISPGNVVRSGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGCCLAFVTVISLLVYKKHKEYNPISPGNVVRSGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLNQKQFEGVS 560
DB 541 PFGNQEKDPLLNQKQFEGVS 560

RESULT 13
US-10-775-972-225
; Sequence 225, Application US/10775972
; Publication No. US20040235072A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C21
; CURRENT APPLICATION NUMBER: US/10/775,972
; CURRENT FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
```



```
; PRIOR APPLICATION NUMBER: 09/685,696
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: 09/662,786
; PRIOR FILING DATE: 2000-09-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-922-124-225

Query Match      100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MECLYFGLGFLLAARLPDAAKRFHDVLGNRPAGYMRHNLQNGWSSDENDWNEKLYP 60
Db      1 MECLYFGLGFLLAARLPDAAKRFHDVLGNRPAGYMRHNLQNGWSSDENDWNEKLYP 60

QY      61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
Db      61 VVKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120

QY      121 RNEAGLSADPYVNTWTANSESDGNGTQSHHNVPDGGPPHHPGWRWNFIYVFTL 180
Db      121 RNEAGLSADPYVNTWTANSESDGNGTQSHHNVPDGGPPHHPGWRWNFIYVFTL 180

QY      181 GQYFQKLGRCVRSVNTANVTLGQLMETVYRHRGRAYVPIAQVKVYVVTDOI PVFV 240
Db      181 GQYFQKLGRCVRSVNTANVTLGQLMETVYRHRGRAYVPIAQVKVYVVTDOI PVFV 240

QY      241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPHFNLNSTINYKWSFGDNTGLEFVSTNHTVN 300
Db      241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPHFNLNSTINYKWSFGDNTGLEFVSTNHTVN 300

QY      301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360
Db      301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

QY      361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db      361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420

QY      421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPDRDPAS 480
Db      421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPDRDPAS 480

QY      481 PLRMANSALISVGLAIFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
Db      481 PLRMANSALISVGLAIFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540

QY      541 PFGNQEKDPLLNQBFKGV 560
Db      541 PFGNQEKDPLLNQBFKGV 560
```

Search completed: December 24, 2005, 04:49:56
Job time : 166 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:36:36 ; Search time 13 Seconds
(without alignments)
307.260 Million cell updates/sec

Title: US-09-853-880A-17
Perfect score: 3026
Sequence: 1 MECLYFLGLLLAARPLD.....FPGNQKDPILKNQEPKGS 560

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
3: /cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
4: /cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
5: /cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
6: /cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
7: /cgm2_6/ptodata/2/pubpaa/US11_NEW_PUB.pap.*
8: /cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	US-10-623-155-225	Sequence 225, App
2	3026	100.0	563	US-10-821-234-1067	Sequence 1067, Ap
3	1029	34.0	206	US-11-090-439-44	Sequence 44, Appl
4	622.5	20.6	662	US-11-090-439-9	Sequence 9, Appl
5	107.5	3.6	564	US-11-022-289-10	Sequence 10, Appl
6	105.5	3.5	427	US-10-995-561-816	Sequence 816, App
7	105.5	3.5	427	US-10-995-561-817	Sequence 817, App
8	105.5	3.5	452	US-10-995-561-818	Sequence 818, App
9	97	3.2	1717	US-11-192-967-2	Sequence 2, Appl
10	97	3.2	1717	US-11-193-715-2	Sequence 2, Appl
11	96.5	3.2	583	US-11-080-991-64	Sequence 64, Appl
12	94.5	3.1	557	US-11-022-289-6	Sequence 6, Appl
13	92	3.0	432	US-11-140-417-4	Sequence 4, Appl
14	92	3.0	435	US-10-510-386-62	Sequence 62, Appl
15	90.5	3.0	551	US-11-022-289-7	Sequence 7, Appl
16	90.5	3.0	557	US-11-022-289-2	Sequence 2, Appl
17	90	3.0	610	US-10-606-302-1	Sequence 1, Appl
18	89.5	3.0	438	US-11-140-417-2	Sequence 2, Appl
19	89.5	3.0	551	US-11-022-289-8	Sequence 8, Appl
20	87.5	2.9	557	US-11-022-289-5	Sequence 5, Appl
21	86.5	2.9	743	US-10-485-517-351	Sequence 351, App
22	86.5	2.9	877	US-10-485-517-200	Sequence 200, App
23	86	2.8	626	US-11-010-748A-1	Sequence 1, Appl
24	85.5	2.8	1390	US-11-063-343-35	Sequence 35, Appl
25	84.5	2.8	175	US-10-821-234-1074	Sequence 1074, Ap

ALIGNMENTS

RESULT 1

US-10-623-155-225
; Sequence 225, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Gary W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C20
; CURRENT APPLICATION NUMBER: US/10/623,155
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-623-155-225

Query Match 100.0%; Score 3026; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 4.2e-255; Indels 0; Gaps 0;
Matches 560; Conservative 0; Mismatches 0;

QY	1	MECLYFLGLLLAARPLDAAKRFHDVLGNERPSAYMREHNQNGSSDENDWNEKLYP	60
Db	1	MECLYFLGLLLAARPLDAAKRFHDVLGNERPSAYMREHNQNGSSDENDWNEKLYP	60
QY	61	VVKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC	120
Db	61	VVKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC	120
QY	121	RNEAGLSADPYVYNNWTAMSESDGNGTGHSHNVFPDGKPPHPPHGRWRNFIYVFHTL	180
Db	121	RNEAGLSADPYVYNNWTAMSESDGNGTGHSHNVFPDGKPPHPPHGRWRNFIYVFHTL	180
QY	181	GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVYVTDQIPVF	240
Db	181	GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVYVTDQIPVF	240
QY	241	TMFQKNDNRSDETFKDLPIPMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN	300
Db	241	TMFQKNDNRSDETFKDLPIPMFDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN	300
QY	301	HTYVLNGTFSNLVTYKAAAGPCPPPPPPPSKPTPSIGPAGDNPLESRLSRIDENCQIN	360
Db	301	HTYVLNGTFSNLVTYKAAAGPCPPPPPPPSKPTPSIGPAGDNPLESRLSRIDENCQIN	360

Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTSLGPAGDNPLLSRIPDENQCIN 360
Qy 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPBESSLIDFVVTQCGSIPTEVCTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPBESSLIDFVVTQCGSIPTEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDRDPAS 480
Db 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDRDPAS 480
Qy 481 PLRMANALISVGCLAIPTVTVISLLVYKHKHYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRMANALISVGCLAIPTVTVISLLVYKHKHYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNQOEKFGVS 560
Db 541 FPGNQEKDPLLNQOEKFGVS 560

RESULT 2
US-10-821-234-1067
; Sequence 1067, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1067
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1067

Query Match 100.0%; Score 3026; DB 6; Length 563;
Best Local Similarity 100.0%; Pred. No. 4.2e-255;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERPSAYMREHNQNLNGWSSDENDWNEKLYP 60
Db 4 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERPSAYMREHNQNLNGWSSDENDWNEKLYP 63
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYEKNC 120
Db 64 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYEKNC 123
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTGSHHNVFPDGKPPFHPGWRWNFIYVPHTL 180
Db 124 RNEAGLSADPYVYNWTAWSESDSGENGTGSHHNVFPDGKPPFHPGWRWNFIYVPHTL 183
Qy 181 GOYFQKLGRCSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQVDKVVVTDQIPVFV 240
Db 184 GOYFQKLGRCSVRVSVNTANTVLGPQLMEVTVYRRHGRAYVPIAQVDKVVVTDQIPVFV 243
Qy 241 TMOKNDNRNSDDEFLKDLPTMFVDVLTDPHSPLNYSTINVKWSFGDNTGLFVSTNHTVN 300
Db 244 TMOKNDNRNSDDEFLKDLPTMFVDVLTDPHSPLNYSTINVKWSFGDNTGLFVSTNHTVN 303
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTSLGPAGDNPLLSRIPDENQCIN 360
Db 304 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTSLGPAGDNPLLSRIPDENQCIN 363
Qy 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPBESSLIDFVVTQCGSIPTEVCTIISDPT 420
Db 364 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPBESSLIDFVVTQCGSIPTEVCTIISDPT 423

Qy 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDRDPAS 480
Db 424 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDRDPAS 483
Qy 481 PLRMANALISVGCLAIPTVTVISLLVYKHKHYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db 484 PLRMANALISVGCLAIPTVTVISLLVYKHKHYNPIENSPGNVVRSGLSVFLNRAKAVF 543
Qy 541 FPGNQEKDPLLNQOEKFGVS 560
Db 544 FPGNQEKDPLLNQOEKFGVS 563

RESULT 3
US-11-090-439-44
; Sequence 44, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberous Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-44

Query Match 34.0%; Score 1029; DB 7; Length 206;
Best Local Similarity 98.4%; Pred. No. 2.1e-82;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERPSAYMREHNQNLNGWSSDENDWNEKLYP 60
Db 1 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERPSAYMREHNQNLNGWSSDENDWNEKLYP 60
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYEKNC 120
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIVYEKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTGSHHNVFPDGKPPFHPGWRWNFIYVPHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTGSHHNVFPDGKPPFHPGWRWNFIYVPHTL 180
Qy 181 GOYFQ 185
Db 181 GLLQ 185

RESULT 4
US-11-090-439-9
; Sequence 9, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberous Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1


```
RESULT 7
US-10-995-561-818
; Sequence 818, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 818
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-818

Query Match      3.5%; Score 105.5; DB 6; Length 427;
Best Local Similarity 32.4%; Pred. No. 0.072;
Matches 36; Conservative 14; Mismatches 38; Indels 23; Gaps 6;

Qy 317 AAAPGCPPPPPPPRPSKPT--PSLGPAGDNPLSLRIPDENCQINRYGH-----FOAT 368
Db 161 AVAPEPCQPLRSPSLDNPTFPNLGSPS-ENPLKRLVPGGEWEFEVTA FYRGQVFQQT 219

Qy 369 ITIVEGILEVNIQMTDVLMP-----VWPPESSLID-----FVTCQG 406
Db 220 ISCEGLRLVG-SEVGDRTLGWPVTLDPDQMSLTDGVMYSYVRHVLSCLG 269

RESULT 8
US-10-995-561-817
; Sequence 817, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 817
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-817

Query Match      3.5%; Score 105.5; DB 6; Length 452;
Best Local Similarity 32.4%; Pred. No. 0.078;
Matches 36; Conservative 14; Mismatches 38; Indels 23; Gaps 6;

Qy 317 AAAPGCPPPPPPPRPSKPT--PSLGPAGDNPLSLRIPDENCQINRYGH-----FOAT 368
Db 161 AVAPEPCQPLRSPSLDNPTFPNLGSPS-ENPLKRLVPGGEWEFEVTA FYRGQVFQQT 219

Qy 369 ITIVEGILEVNIQMTDVLMP-----VWPPESSLID-----FVTCQG 406
Db 220 ISCEGLRLVG-SEVGDRTLGWPVTLDPDQMSLTDGVMYSYVRHVLSCLG 269

RESULT 9
US-11-192-967-2
; Sequence 2, Application US/11192967
; Publication No. US20050262587A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
```

```
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; TITLE OF INVENTION: Lepidopteran Insects and Methods of Use
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/192,967
; CURRENT FILING DATE: 2005-07-29
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-11-192-967-2

Query Match      3.2%; Score 97; DB 7; Length 1717;
Best Local Similarity 21.2%; Pred. No. 2.5;
Matches 91; Conservative 67; Mismatches 156; Indels 116; Gaps 26;

Qy 154 NVPPDGKPPPH-----PGWRMNPFIY--VFHTLQY-----FQKLGRCVRSVNTANVTL 203
Db 434 DVFPFGAASAFYIAPGSGYQRTPIMGITINHMLDYEDVIFQ-----NIIKVKAVDMNN 488

Qy 204 GPQLMEVTVYRRHGRAVPIAQKDVVV--TQIIPVFTVMFQKNDNRSSDETEFLKDLPI 261
Db 489 ASHVGEALVY-----VNLIWNDELPIFE-----ESSYSASFKETVGAGFPV 530

Qy 262 -----MFDVLIHDPHFLNYSTINYKWSFGDNTG-LFVSTN-----HTVNHTYVL- 305
Db 531 ATVLALDRDIDVVVHS-----LGNADVLYL-F-IDESTGEIFVSMDDAFYHRQNTL FQV 585

Qy 306 -----NGTFSMLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLR----- 351
Db 586 RADDTLGDGPHNTVTTLQVIELEDVNNTPPTLRLPR---STPSVEENVVPEGYEISREITA 642

Qy 352 -IPDENCQINRYGHFOATITIVEGILEVNIQMTDVLMPVWPPESSLIDFVTCQSGIPT 410
Db 643 TDPDTSAYLWFEDWDSTWATKQG-RETN-----PTEYVGCIVIEIYIYFTEGNGRS 692

Qy 411 EVCTIISDPTCEITQNTVCSFVDVD--EMCLLTVRTFNGSGTYCVNLTGLGDDTSLA-LT 467
Db 693 AIGRLVVQ---EIRDNT---IDFEFEMLYLTVR-----VRDLNTVIGDDVDEATFT 739

Qy 468 STLISVDPDRDPASPLRMANSALISV-----GCL--AIFVTVISLLVYKKHK-EYNP 515
Db 740 ITIIDMNDNAPI-----FANGTLTQTMVRVBLAASGTLIGSVLATDIDGPLYNQVRYTIQ 795

Qy 516 IENSPGNVVR 525
Db 796 RNNTPEGLVK 805

RESULT 10
US-11-193-715-2
; Sequence 2, Application US/11193715
; Publication No. US20050260675A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; TITLE OF INVENTION: Lepidopteran Insects and Methods of Use
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/193,715
; CURRENT FILING DATE: 2005-07-29
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
```

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Oestrinia nubialis
US-11-193-715-2

Query Match      3.2%; Score 97; DB 7; Length 1717;
Best Local Similarity 21.2%; Pred. No. 2.5;
Matches 91; Conservative 67; Mismatches 156; Indels 116; Gaps 26;

QY 154 NVPPGKPPPH-----PGRWRNRYI--VFHTLQY-----FQKGRCSVRVSVNTANVTL 203
DB 434 DVFFPGAASAFIAPGSGYQRTFIMGTINGHTMLDYEDVIFQ-----NIIIKVKAVDNN 488
QY 204 GPQLMEVTVYRRHGRAYVPIAQVKDVVVV--TDQIPVFTVFPQKNDRSSDETFLKDLPI 261
DB 489 ASHVGELVY-----VNLINWDELPIFE---ESSYSASFKEITVGAGFPV 530
QY 262 -----MFVLIHDPHSHFLNYSTINYKWSFGDNTG-LFVSTN-----HTVNHYYVL- 305
DB 531 ATVALDRDIDVVVHS-----LMGNAVDFLF-IDESTGEIFVSMDDAFDYHRQNTLFFQV 585
QY 306 -----NGTFSNLITVRAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSR----- 351
DB 586 RADDTLGDPGPHNTVTTLVIELEDVNNPTTLRLPR-----STPSVEENVPBGYEISREITA 642
QY 352 -IPDENCOINRYGHFOATITIVEGILEVNIQMOTDVLMPVPWPESSLIDFVVTCGSIPT 410
DB 643 TDPDTSAVLNPEIDWDSWATKQG-REYN-----PTEVGCIVIETIYPTGNGRS 692
QY 411 EVCTIISDPTCEITONTVCSPVDVD--EMCLLTVRTFTNGSGTYCVNLTLGDDTSLA-LT 467
DB 693 AIGRLVWQ---EIRDNV---IDPEFEFMYLTVR-----VRDLNTVIGDDYDEATFT 739
QY 468 STLISVPDRDPASPLRMANSALISV-----CCL--AIPVTVISLLVYKHKH-EYNP 515
DB 740 ITIIMDNAPT-----FANGTLTQTMVRVRELAASGTLIGSVLATDIDGLYNQVRYTIQP 795
QY 516 IENSPGNVVR 525
DB 796 RNNTEGLVK 805

RESULT 11
US-11-080-991-64
; Sequence 64, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; PRIOR FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-080-991-64

Query Match      3.2%; Score 96.5; DB 7; Length 583;
Best Local Similarity 19.4%; Pred. No. 0.66;
Matches 80; Conservative 53; Mismatches 137; Indels 143; Gaps 16;

QY 205 POLMEVTVYRRHGRAYVPIAQ-VKDVVVVTDQIPVFTVFPQKNDRSSDETFLKDL----- 259

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Oestrinia nubialis
US-11-193-715-2

Query Match      3.1%; Score 94.5; DB 7; Length 557;
Best Local Similarity 23.1%; Pred. No. 0.92;
Matches 74; Conservative 40; Mismatches 135; Indels 71; Gaps 19;

QY 86 SPALVGSNITFAVNLIPRCOKEDANGNIVYEKKNCRNEAGLSADPYV--YNTAWSESDG 144
DB 114 APELLGGSVF-----LFPKPKDTLMISRTPEVTCVVVDVSHEDDEVKFNVTV----- 162
QY 145 ENGTCQSHHNVFPDGKPPPHHFGWRWNFIYVFH---TLGOYFQKLGRCSVRSVYNTANV 201
DB 163 ---DGVEVHNAKTPREBQYNSYTVVSVLTVLHQDLWLNKKEY-----KCKV-----SNK 209
QY 202 TLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQI---PVFTVFPQKN-----DR 248
DB 210 AL-PAPIEKTISKAGQPRP- -QVYTLPPSRDELTKQVSLTCLVKGFFYPSDIAVRES 266
QY 249 NSSDETFLKDLPIFMDVLIHDPHSHFLNYSTINY---KWSFGDNTGLF-VSTNHTVNHVTV 304
DB 267 NGQPENNYKTTFP-----PVLDSGGSFFL-YSKUTVVKSRWQOQN---VFSCSVNHEALHNY 319
QY 305 LNGTFSNLITVRAAAGP-----PCPPPPPPRPSKPTPSL-GPAGDNPLELSRIPDENCQ 358
DB 320 TQKSLSL-----SPGDKDTHTCPPSPAPELLGGPSVFLPPKPKDTLMISRTPEVTCV 372

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Oestrinia nubialis
US-11-193-715-2

Query Match      3.1%; Score 94.5; DB 7; Length 557;
Best Local Similarity 23.1%; Pred. No. 0.92;
Matches 74; Conservative 40; Mismatches 135; Indels 71; Gaps 19;

QY 86 SPALVGSNITFAVNLIPRCOKEDANGNIVYEKKNCRNEAGLSADPYV--YNTAWSESDG 144
DB 114 APELLGGSVF-----LFPKPKDTLMISRTPEVTCVVVDVSHEDDEVKFNVTV----- 162
QY 145 ENGTCQSHHNVFPDGKPPPHHFGWRWNFIYVFH---TLGOYFQKLGRCSVRSVYNTANV 201
DB 163 ---DGVEVHNAKTPREBQYNSYTVVSVLTVLHQDLWLNKKEY-----KCKV-----SNK 209
QY 202 TLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQI---PVFTVFPQKN-----DR 248
DB 210 AL-PAPIEKTISKAGQPRP- -QVYTLPPSRDELTKQVSLTCLVKGFFYPSDIAVRES 266
QY 249 NSSDETFLKDLPIFMDVLIHDPHSHFLNYSTINY---KWSFGDNTGLF-VSTNHTVNHVTV 304
DB 267 NGQPENNYKTTFP-----PVLDSGGSFFL-YSKUTVVKSRWQOQN---VFSCSVNHEALHNY 319
QY 305 LNGTFSNLITVRAAAGP-----PCPPPPPPRPSKPTPSL-GPAGDNPLELSRIPDENCQ 358
DB 320 TQKSLSL-----SPGDKDTHTCPPSPAPELLGGPSVFLPPKPKDTLMISRTPEVTCV 372
```


Db 210 AL-PAPIEKTISKAKQPREP--QVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWES 266
 QY 249 NSSDETFLKOLPIMFDVLIHDPHFPLNYSTINY---KWSFGDNTGLP-VSTNHTVNNHTYV 304
 Db 267 NGQPENNYKTPE---PVLDSGSGFEL-YSKLTVDKSRWQOQN---VFSCSVMHREALHNY 319
 QY 305 LNGTESLNLTVKAAAPG--PCPPP-----PPPRPSKPTPSLGPAGDNELELSR 351
 Db 320 TOKSLSL-----SPGKPPCPAPELLGGSVFLFPKPK-----DTLMISR 359
 QY 352 IPDENCQINRYGHFOATITI---VEGI 375
 Db 360 TPEVTCVVVDVSHEDPEVKFNWYVDGV 386

Search completed: December 24, 2005, 04:50:19
 Job time : 20 secs

THIS PAGE BLANK (USPTO)